



TECHNICKÁ UNIVERZITA V LIBERCI
Fakulta mechatroniky, informatiky
a mezioborových studií ■

WEBOVÁ APLIKACE PRO SPRÁVU SEO

Bakalářská práce

Studijní program: B2646 – Informační technologie
Studijní obor: 1802R007 – Informační technologie
Autor práce: **Lukáš Javůrek**
Vedoucí práce: Ing. Petr Kretschmer





TECHNICAL UNIVERSITY OF LIBEREC
Faculty of Mechatronics, Informatics
and Interdisciplinary Studies ■

WEB APPLICATION FOR MANAGING SEO

Bachelor thesis

Study programme: B2646 – Information Technology
Study branch: 1802R007 – Information Technology
Author: **Lukáš Javůrek**
Supervisor: Ing. Petr Kretschmer



ZADÁNÍ BAKALÁŘSKÉ PRÁCE

(PROJEKTU, UMĚLECKÉHO DÍLA, UMĚLECKÉHO VÝKONU)

Jméno a příjmení: **Lukáš Javůrek**
Osobní číslo: **M12000139**
Studijní program: **B2646 Informační technologie**
Studijní obor: **Informační technologie**
Název tématu: **Webová aplikace pro správu SEO**
Zadávací katedra: **Ústav nových technologií a aplikované informatiky**

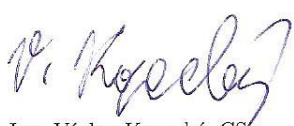
Z á s a d y p r o v y p r a c o v á n í :

1. Seznamte se s metodami SEO analýzy.
2. Sestavte na základě obecných a zveřejněných doporučení vyhledávačů metodiku vyhodnocování SEO, kterou použijete v aplikaci.
3. Navrhněte strukturu aplikace a databáze.
4. Realizujte vlastní aplikaci vhodnými vývojovými prostředky a ověřte její funkčnost.
5. Vytvořte instalační distribuci a stručný návod k použití aplikace.
6. Zhodnoťte účinnost SEO po aplikaci vydaných doporučení na sledované stránky.

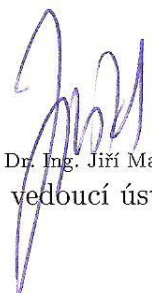
Rozsah grafických prací: **dle potřeby**
Rozsah pracovní zprávy: **cca 40 stran**
Forma zpracování bakalářské práce: **tištěná/elektronická**
Seznam odborné literatury:

- [1] Gilmor, J. W., Velká kniha PHP a MySQL 5 - kompendium znalostí pro začátečníky i profesionály, Zoner Press, 2007
[2] Kosek, J., PHP - tvorba interaktivních internetových aplikací, Grada Publishing, 1999, ISBN 80-7169-373-1

Vedoucí bakalářské práce: **Ing. Petr Kretschmer**
Ústav nových technologií a aplikované informatiky
Konzultant bakalářské práce: **Ing. Vojtěch Wrnata**
Ústav nových technologií a aplikované informatiky
Datum zadání bakalářské práce: **20. října 2014**
Termín odevzdání bakalářské práce: **15. května 2015**


prof. Ing. Václav Kopecký, CSc.
děkan




prof. Dr. Ing. Jiří Maryška, CSc.
vedoucí ústavu

V Liberci dne 20. října 2014

Prohlášení

Byl jsem seznámen s tím, že na mou bakalářskou práci se plně vztahuje zákon č. 121/2000 Sb., o právu autorském, zejména § 60 – školní dílo.

Beru na vědomí, že Technická univerzita v Liberci (TUL) nezasahuje do mých autorských práv užitím mé bakalářské práce pro vnitřní potřebu TUL.

Užiji-li bakalářskou práci nebo poskytnu-li licenci k jejímu využití, jsem si vědom povinnosti informovat o této skutečnosti TUL; v tomto případě má TUL právo ode mne požadovat úhradu nákladů, které vynaložila na vytvoření díla, až do jejich skutečné výše.

Bakalářskou práci jsem vypracoval samostatně s použitím uvedené literatury a na základě konzultací s vedoucím mé bakalářské práce a konzultantem.

Současně čestně prohlašuji, že tištěná verze práce se shoduje s elektronickou verzí, vloženou do IS STAG.

Datum:

Podpis:

Poděkování

Děkuji vedoucímu práce Ing. Petru Kretschmerovi za metodické vedení a cenné rady během tvorby této práce, také děkuji panu Mgr. Jiřímu Vranému Ph.D., který mi poskytl konzultaci, podporu a zajímavé myšlenky k řešené problematice a všem dalším, kteří přidali cenné rady k této práci.

Děkuji především své rodině, že mi byla oporou, podporovala mě a po celou dobu mého studia mi poskytovala klidné zázemí ke studiu.

Abstrakt

Tato bakalářská práce se zabývá optimalizací stránek pro internetové vyhledávače, této problematice se také říká SEO. Prvním cílem této práce bylo nastudování principů a důležitých faktorů ohledně SEO problematiky. Druhým cílem práce bylo vytvoření nástroje v podobě online aplikace pro analyzování internetových stránek s výstupy, které obsahují rady, tipy a obecné informace pro zlepšení jednotlivých testovaných částí webové stránky. Přínosem práce je online nástroj se snadnou instalací a velkým počtem testů.

Nástroj je vytvořen jako serverová aplikace za pomoci skriptovacího jazyka PHP a nástroje Nette Framework. K výsledné aplikaci lze přistupovat z libovolného přístroje s přístupem na internet, díky použití responzivního designu, který využívá Twitter Bootstrap.

Klíčová slova:

PHP, Nette Framework, SEO, webová aplikace, nástroj pro analýzu webových stránek

Abstract

This bachelor thesis is about optimizing websites for internet browsers. This issue is also called SEO. The first aim of this work was studying principals and important factors of SEO. The second aim was to create a tool in form of online application for analyzing websites with outputs including advices, tips and general information for the improvement of the various tested parts of the site. The asset of this thesis is the online tool with easy installation and big number of tests.

The tool is made as a server application by PHP script language and tool called Nette Framework. The resulting application can be used from any device with internet access, thanks to responsive design, which uses the Twitter Bootstrap.

Keywords:

PHP, Nette Framework, SEO, web application, tool for website analysis

Obsah

1	ÚVOD.....	11
2	TEORETICKÁ ČÁST.....	13
2.1	INTERNETOVÉ KATALOGY A FULLTEXTOVÉ VYHLEDÁVAČE.....	13
2.1.1	Internetové katalogy.....	13
2.1.2	Fulltextové vyhledávače.....	13
2.1.3	Optimalizace pro vyhledávače.....	15
2.2	ON-PAGE FAKTORY.....	16
2.2.1	Technické překážky indexace.....	16
2.2.2	Technické faktory.....	18
2.3	OFF-PAGE FAKTORY.....	26
2.3.1	Zpětné odkazy.....	26
2.3.2	Ranky.....	28
2.3.3	PageRank.....	28
2.3.4	S-Rank.....	28
2.3.5	Alexa Rank.....	29
2.4	ZAKÁZANÉ TECHNIKY.....	29
2.4.1	Skryté texty.....	30
2.4.2	Skryté odkazy.....	30
2.4.3	Cloaking.....	30
2.4.4	Doorway (gateway).....	30
2.4.5	Komentářový spam.....	31
2.4.6	Link farmy.....	31
2.4.7	Tajné přesměrování.....	31
2.4.8	Automatické registrace do katalogů.....	31
2.4.9	Manipulace s rank hodnotou.....	31
2.4.10	Microsite.....	32
2.5	NETTE FRAMEWORK.....	32
2.5.1	Výhody.....	32
2.5.2	Požadavky.....	33
2.5.3	Stahování a instalace.....	33
2.5.4	První spuštění.....	33
2.5.5	MVC aplikace.....	34
2.5.6	Ladící nástroj Tracy.....	35
2.5.7	Šablonovací jazyk Latte.....	36
2.5.8	Databáze.....	36
2.6	TWITTER BOOTSTRAP.....	37
2.7	MYSQL.....	37

2.7.1	<i>MyISAM</i>	38
2.7.2	<i>InnoDB</i>	38
3	CÍL PRÁCE	39
4	POSTUP VÝVOJE APLIKACE	40
4.1	POUŽITÉ TECHNOLOGIE	40
4.2	KONTROLA MINIMÁLNÍCH POŽADAVKŮ	40
4.3	NÁVRH DATABÁZE	41
4.4	POSTUP VYTVÁŘENÍ TESTŮ.....	43
4.5	ON-PAGE FAKTORY	44
4.5.1	<i>Získání zdrojového kódu</i>	44
4.5.2	<i>Doctype</i>	45
4.5.3	<i>Hlavička</i>	45
4.5.4	<i>Kódování stránky</i>	45
4.5.5	<i>Titulek</i>	45
4.5.6	<i>Metainformace</i>	46
4.5.7	<i>Popisek</i>	46
4.5.8	<i>Klíčová slova</i>	46
4.5.9	<i>Autor</i>	46
4.5.10	<i>Informace pro roboty</i>	47
4.5.11	<i>Nadpisy</i>	47
4.5.12	<i>Odkazy</i>	47
4.5.13	<i>Návrh klíčových slov</i>	48
4.5.14	<i>Kontrola klíčových slov v URL</i>	48
4.5.15	<i>Alternativní texty u obrázků</i>	48
4.5.16	<i>Délky</i>	48
4.5.17	<i>Čas načtení zdrojového kódu</i>	49
4.5.18	<i>robots.txt</i>	49
4.5.19	<i>sitemap.xml</i>	49
4.5.20	<i>Vnořené tabulky</i>	49
4.5.21	<i>Průchod webem</i>	50
4.6	DALŠÍ FUNKCE SEO APLIKACE.....	50
4.6.1	<i>Google</i>	51
4.6.2	<i>Bing</i>	51
4.6.3	<i>Seznam</i>	51
4.6.4	<i>Zpětné odkazy</i>	51
4.6.5	<i>Pozice ve vyhledávačích</i>	52
4.6.6	<i>Ranky</i>	52
4.6.7	<i>Analýza validity</i>	52
4.7	JEDNOTNÁ INFORMAČNÍ VRSTVA.....	52

4.8	REGISTROVANÁ ČÁST	52
4.8.1	<i>Správa aplikací</i>	54
4.8.2	<i>Diskuze</i>	56
5	VYHODNOCENÍ ŘEŠENÍ.....	58
5.1	ČASOVÁ A PAMĚŤOVÁ NÁROČNOST.....	58
5.2	ZLEPŠENÍ POZIC TESTOVANÉ WEBOVÉ APLIKACE	59
5.3	SROVNÁNÍ EXISTUJÍCÍCH APLIKACÍ	60
5.3.1	<i>Seo-servis</i>	60
5.3.2	<i>Seonastroje</i>	60
5.3.3	<i>Otestujte Web</i>	61
5.4	MOŽNOSTI ROZŠÍŘENÍ	62
6	ZÁVĚR	63
	POUŽITÉ ZDROJE	65

Seznam obrázků

Obrázek 1: Původní architektura Google vyhledávače	14
Obrázek 2: Výsledky vyhledávače	21
Obrázek 3: Algoritmus pro výpočet PageRanku	28
Obrázek 4: Struktura adresářů Sandboxu	33
Obrázek 5: Použití modulů v aplikaci	34
Obrázek 6: Tracy panel.....	35
Obrázek 7: Nette identita	36
Obrázek 8: Nette - podporované databáze.....	37
Obrázek 9: Diagram databáze	43
Obrázek 10: Načtení zdrojového kódu	44
Obrázek 11: Registrace.....	53
Obrázek 12: Přihlášení.....	53
Obrázek 13: Správa aplikací.....	54
Obrázek 14: Vybraná webová aplikace	55

Obrázek 15: Nastavení tématu diskuze	55
Obrázek 16: Miniaturní diskuze	56
Obrázek 17: Diskuze	57
Obrázek 18: Chat.....	57

Seznam tabulek

Tabulka 1: Velikost a čas načtení webových stránek.....	58
Tabulka 2: Časová a paměťová náročnost - domácí počítač.....	59
Tabulka 3: Časová a paměťová náročnost - server hosting.....	59
Tabulka 4: Srovnání rychlosti testu na zdrojovém kódu.....	62

Seznam zkratek

AJAJ	Asynchronous JavaScript and JSON
AJAX	Asynchronous JavaScript and XML
API	Application Programming Interface
ASCII	American Standard Code for Information Interchange
BAN	Boundary Access Nodes
CSRF	Cross-site Request Forgery,
CSS	Cascading Style Sheets
GNU GPL	GNU General Public License
HTML	Hypertext Markup Language
HTTP	Hypertext Transfer Protocol
MVC	Model View Controller
MySQL	My Structured Query Language
New BSD	New Berkeley Software Distribution
PDO	PHP Data Object
PHP	Hypertext Preprocessor
PPC	Pay Per Click
SEO	Search Engine Optimization
SPAM	spiced ham
SQL	Structured Query Language
URL	Unique Resource Locator
XSS	Cross Site Script

1 Úvod

Prvotní myšlenka byla vytvořit komplexní autonomní nástroj pro kontrolu a testování všech dílčích prvků ovlivňující pořadí ve výsledcích vyhledávačů, tento nástroj měl testovat jednotlivé faktory SEO optimalizace a na základě toho vytvořit rady a tipy pro zlepšení a registrovat testovaný web v různých internetových katalozích.

Jak jistě všichni víme, tak Internet roste a mění se každým dnem, zanikají a vznikají firmy s jejich vnitřními sítěmi, také webové stránky vznikají a zanikají. Internet se stává stále důležitější součástí života dnešní populace. A stále méně uživatelů hledá nějaké informace, místo toho hledají zábavu, odreagování, data, zboží a v neposlední řadě sociální sítě.

V dnešní době si téměř každý člověk může vytvořit nějakou vlastní webovou stránku. A není to vůbec složité. Nejjednodušší je podle mne vytvořit blog, další úrovní jsou redakční systémy. Poslední možnost bych rozdělil na dvě kategorie podle potřebného času na vytvoření webové stránky. A to na frameworky, které nám ulehčí mnoho práce, a poslední je možnost, kdy si programátor napíše celou webovou stránku sám. Ovšem proč si trochu neulehčit práci.

A právě kvůli tomuto obrovskému množství webových stránek dnes už nestačí mít pouze nějakou stránku na Internetu. Stále důležitějším je mít kvalitní obsah. Také více nabývá na důležitosti grafický vzhled webové stránky, kde se stále více uživatelů přiklání k jednoduchému vzhledu, bez všelijakých „blikajících“ částí a nepříjemně barevných kombinací, u kterých Vás začnou bolet oči hned po načtení. V dnešní době každá firma, která nemá jedinečný sortiment, chce být na předních pozicích ve vyhledávacích kvůli konkurenci, která je opravdu velká...

Optimalizace pro vyhledávače je důležitá, neboť je dokázáno, že většina uživatelů při vyhledání pomocí vyhledávačů věnuje pozornost pouze první stránce s výsledky a většina z většiny pouze prvním třem výsledkům. Na druhou stránku se už téměř nikdo z nich nepodívá.

Na základě těchto zkušeností vznikl obor SEO. Mnoho článků na toto téma se rozchází v důležitosti jednotlivých optimalizačních metod a optimalizovaných prvků. Proto bych opomenul důležitost, která se velice často uvádí. Za poslední tři

roky se velice rozšířily webové stránky, které Vám slibují přední pozice ve výsledcích při vyhledávání ve vyhledávačích, některé píší o tom, že se dostanete na přední pozice pro konkrétní klíčová slova, někde si tato slova (fráze) můžete zvolit sami, jinde je volí jejich systém z Vašich klíčových slov a popřípadě i nadpisů úrovně jedna. Většina těchto stránek není moc důvěryhodná, protože jejich vlastníci se snaží na tyto informace nalákat vlastníky webových stránek, kteří někde zaslechli o optimalizaci, ale nic víc o ní neví. Marketing je jejich hlavní priorita. SEO optimalizace se dnes stále více stává zdrojem obživy. Proto velké množství těchto organizací používá bodové ohodnocení, kde po dosažení maximálního počtu bodů Vám slibují přední pozice. Ti méně nečestní se zmíní o tom, že se na přední pozice nemusíte dostat hned. Zase na druhou stranu, pokud se opravdu chcete dostat na přední pozice a nechcete strávit desítky až stovky hodin, tak nějakou firmu budete nejspíše potřebovat.

I když Vaši webovou stránku dokáží dostat na přední místa, už Vám nemohou zaručit, že tam také zůstanete na pořád, ani na jak dlouho tam budete. A to díky tomu, že zase jinde mohl někdo udělat SEO optimalizaci na jiné webové stránce a tím Vás posunout dále.

2 Teoretická část

2.1 Internetové katalogy a fulltextové vyhledávače

2.1.1 Internetové katalogy

Internetový katalog je na Internetu seznam odkazů na webové stránky, které jsou seříděny do stromu kategorií a podkategorií. Internetové stránky do katalogu navrhuji převážně jejich majitelé a editoři katalogu tyto návrhy kontrolují a zveřejňují, to umožňuje přesné zařazení odkazů do jednotlivých kategorií katalogu.

Katalogy nám umožňují procházet stromovou strukturou podle kategorií a podkategorií, dále prohlížení a vyhledávání odkazů podle určitého slova či frází. A v neposlední řadě i nejzajímavější část pro vlastníky stránek je přidávání odkazů do databáze katalogu.

Katalogy jsou projekty velkých portálových internetových firem, třeba Seznam nebo Centrum. O katalogy je potřeba se starat, nové odkazy na webové stránky musejí projít kontrolou, zda majitel navrhl odkaz do správné kategorie.

V dnešní době ovšem internetové katalogy upadají, jednak kvůli nutnosti lidské kontroly, jelikož lidský čas je stále dražší a společensky cennější a především díky fulltextovým vyhledávačům. Ovšem díky vyhledávačům se nám stále vyplatí registrovat webové stránky v katalogích, díky tomu získáme další důvěryhodné odkazy, díky čemu se posuneme vzhůru ve vyhledávání. [1], [2]

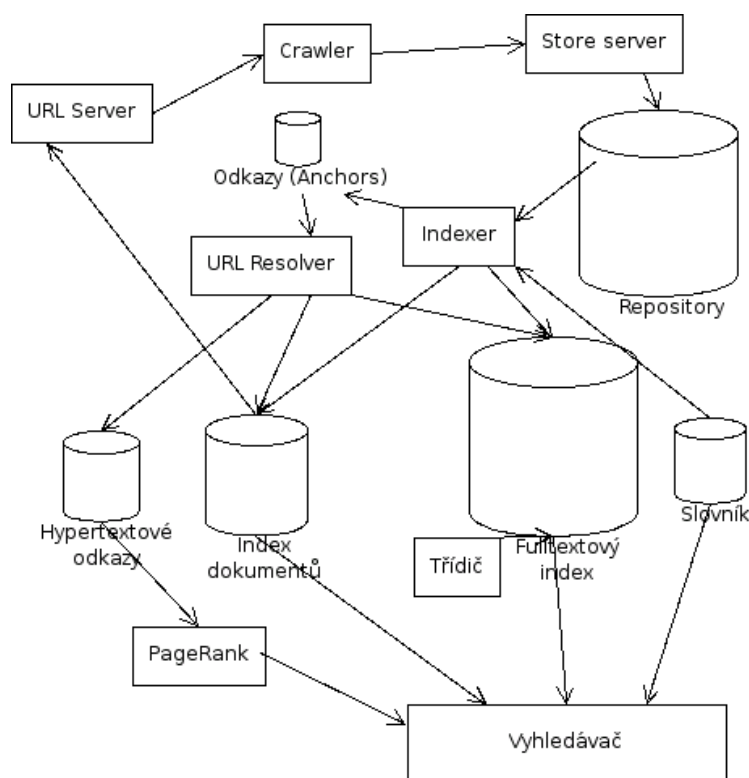
2.1.2 Fulltextové vyhledávače

Fulltextový vyhledávač pracuje na zcela jiném principu než katalog. Můžeme si představit, že je to jakýsi program, který postupně prochází jednotlivé internetové stránky na celém světě a zapisuje je do své databáze (tzv. provádí indexaci do své databáze).

O indexaci stránek se stará sám vyhledávač prostřednictvím robota, který se po Internetu pohybuje samostatně a nelze mu nařídit, kdy se má na Vaši stránku znovu vrátit a zaindexovat případné změny. Můžete pouze robotovy zakázat, aby neindexoval část vašich stránek, nebo aby vaše stránky neindexoval vůbec.

V začátcích fulltextových vyhledávačů nebyly výsledky vyhledávání tak relevantní jako „dnes“. Pokud jsme vyhledávali nějakou frázi (např. autíčko na baterky), tak vyhledávač nám třeba vracel na prvních pozicích ve výsledcích i takové stránky, na kterých se nevyskytovala celá hledaná fráze, ale třeba jen její část (např. autíčko, nebo na baterky). Jako první přišel s touto změnou Google, který se snažil nejdříve zobrazovat webové stránky, na kterých se vyskytovala celá hledaná fráze a až na dalších pozicích byly stránky, které neobsahovaly celou hledanou frázi. Google byl také první, který začal webovým stránkám přiřazovat jejich důležitost. Podle tohoto kritéria postupně zobrazuje nalezené výsledky, čímž jsou relevantnější.

Pro nastínění původní architektury celého vyhledávacího a indexovacího systému je přiložen následující Obrázek 1. [1]



Obrázek 1: Původní architektura Google vyhledávače

Je známo mnoho faktorů ovlivňujících výsledky vyhledávání ve fulltextových vyhledávačích a mezi SEO specialisty existuje velice mnoho rozporuplných názorů především v tom, jakou mají jednotlivé faktory váhu. Jak je tomu doopravdy vědí jen návrháři, popřípadě programátoři vyhledávačů, kteří vše přísně tají. Níže jsou uvedeny ty nejznámější a obecně nejdiskutovanější faktory, které mají podíl na dobré pozici stránek ve výsledcích vyhledávání. [3], [4], [5]

Vyhledávač prohledává v první řadě text stránky, titulek, popisek, klíčová slova a adresu. Zde hledá hledanou frázi (dotaz hledajícího). Záleží na tom, kde se klíčové slovo (hledaný výraz) na stránce nachází a jak je formátováno. Nejvyšší váhu mají klíčová slova v titulku, popisku, klíčových slovech a v nadpisech, kde hlavní nadpis <h1> má větší váhu než ostatní nadpisy a než běžně naformátovaný text na stránce.

Každou stránku si vyhledávač ohodnotí a přidělí mu tzv. skóre kvality. Google používá označení PageRank a Seznam S-rank, který vyjadřuje důležitost stránky.

Důležitý je počet zpětných odkazů vedoucích na vaši stránku a jejich anchor text (příklad: toto je anchor text). Vyhledávače si řadí stránky do tematických skupin. To znamená, že zpětné odkazy na stránkách s podobnou tematikou jsou hodnoceny daleko lépe, než odkazy směřující z webu, který má úplně jiné zaměření než ten váš. Ovšem v horším případě může dojít i k penalizaci za odkazy mířící ze stránek s odlišnou tematikou.

2.1.3 Optimalizace pro vyhledávače

Zkratka SEO pochází z anglického Search Engine Optimization, tedy optimalizace pro vyhledávače. Vyhledávače jako Google, Seznam, Jyxo, či Bing využívají postupů, které určují relevanci webu ve vztahu k vyhledávanému výrazu. Hledáte-li například sousloví „optimalizace pro vyhledávače“, vyhledávač najde ve své databázi všechny stránky, které tuto frázi obsahují, a interně si je ohodnotil. Ty, které mají největší hodnotu, poté zobrazuje jako první a ty s menší hodnotou pod nimi. To je podstata optimalizace pro vyhledávače.

Obecně se jedná o řadu úkonů a operací, jejichž cílem je především, zlepšení pozic ve fulltextových vyhledávačích = zásadní zlepšení návštěvnosti cílovou skupinou zákazníků. SEO optimalizace je dnes základem úspěchu webových stránek. Web, který je málo navštěvován, neplní jednu ze svých zásadních funkcí.

SEO Search Engine Optimization - SEO optimalizace pro vyhledávače je v současné době nejčastější a nejefektivnější způsob zvyšování počtu přístupů

na webové stránky. Vlastně ani nejde o nic nového ani objevného, důvody, proč je nyní optimalizace pro vyhledávače tak populární jsou asi dva. Jedním důvodem je větší využívání fulltextového vyhledávání oproti katalogovému, čímž se zvyšuje efektivita optimalizace, jelikož většina fulltextových vyhledávačů umí najít hledané slovo v různých tvarech a v téměř všech textových částech. Druhým důvodem je pak poměrně nízká investice, která přináší užitek trvalejšího rázu než většina ostatních reklam. Na druhou stranu, pokud natrefíte na špatnou firmu zabývající se SEO, tak platby za PPC reklamy bude Váš nejmenší problém. [6], [7]

2.2 On-page faktory

Optimalizace webových stránek je dlouhodobý proces, který potřebuje pečlivé naplánování. Čas věnovaný úvodní analýze se vám několikanásobně vrátí. Před započítím práce byste měli znát odpovědi na několik zásadních otázek.

- Jaké je zaměření vaší firmy, jaký je cíl stránek?
- Jaká je cílová skupina návštěvníků?
- Kdo je hlavní konkurence, čím se lišíte?
- Je současná doména optimální?
- Jak jsou na tom stránky nyní?
- Pro jaká klíčová slova bude stránka optimalizována?

Velice důležité je nezapomenout jednu ze základních informací o SEO optimalizaci, a to, že optimalizace se dělá pro každou stránku na vašem webu zvlášť. Každá stránka má mít svá konkrétní klíčová slova, popisek, titulek a obsah. Určitá slova se mohou samozřejmě opakovat i v metainformacích, tím může být například název firmy. [8], [9], [10]

2.2.1 Technické překážky indexace

Mezi technické překážky indexace patří obsah, který máme na stránkách, a není textový. To je například **JavaScript, Flash, Java, obrázky a další objekty**. Procházení těchto zdrojových kódů, nebo spouštění je celkem náročné na čas a paměť. Tudíž se s velkou pravděpodobností ztratí mnoho informací, nebo nemusejí být pochopeny tak, jak mají být chápány. [8]

Dalším velkým omezením může být **duplicitní obsah**. Klasickým příkladem může být URL `www.priklad.cz`, `priklad.cz`, `priklad.cz/index.php` a `www.priklad.cz/index.php`. Dříve právě s indexem a přístupem s `www` a bez `www` byl velký problém, ale dnes už vyhledávače umí tento problém řešit sami. Buď si jednu adresu vyberou a z ostatních udělá kanonickou množinu, nebo první navštívenou URL uloží a z ostatních udělá kanonickou množinu. Dalším problémem bývají filtry a řazení, kde nám obsah stránky trochu mění v závislosti na změně v URL a to většinou nějakým parametrem. Zde je nejvhodnější, abyste si sami vybrali jedno URL. A to tu, na které máte nejvíce položek. Ostatním kanonickým stránkám můžete vložit do hlavičky `<link rel="canonical" href="www.priklad.cz/bazeny?filtr=vse">`, nebo `<meta name="robots" content="noindex, nofollow">` pro zakázání indexace a sledování odkazů ze stránky. [11], [12], [13]

Jako další technická překážka může být **validita stránky**. Zpravidla pokud webový browser dokáže stránku vykreslit podle vašich představ, tak by problém být neměl. Ale může se klidně stát, že vyhledávač veškerý text zaindexuje, ale už kvůli chybám v HTML kódu nemusí pochopit přesný význam a důležitost zaindexovaného textu. [8]

Velikost HTML dokumentu zřejmě sama o sobě hodnocena nebude. Ale pokud budou ve zdrojovém HTML kódu další jazyky, které mohou být v samostatných souborech, tak bude zdrojový kód zbytečně velký, déle se generuje a jeho přenos k uživateli trvá déle. Proto by měl být JavaScript a CSS v samostatných souborech, které jsou nalinkovány. Každopádně Google věnuje stále větší pozornost mobilním zařízením a optimalizaci pro ně, v Google Webmasters Tools radí, aby velikost všech souborů byla co nejmenší.

Rychlost odpovědi webové aplikace je také hodnocena a to poměrně velkou váhou. Největší váhu tomuto faktoru přikládá Google. Vyhledávače preferují stránky s rychlou odezvou. A je to celkem logické. Téměř každý uživatel má rád rychlé webové stránky, jelikož šetří drahocenný čas a málo uživatelů bude procházet webovou aplikaci, které trvá sestavení každé stránky například 2 – 4 sekundy. [8], [14]

Vyhledávače při indexaci webu procházejí každou stránku zvlášť, ale také sbírají i informace napříč webovou aplikací. Jednou sbíranou skupinkou jsou i **neplatné odkazy**. Toto se může stát například v e-shopech při ukončení prodeje polož-

ky nebo po smazání článku. Důležité je, aby vyhledávač dostal informaci o neplatnosti odkazu. Nejlepším řešením je vrátit HTTP kód 404. Ovšem tento kód by měl být už v hlavičce odezvy ze serveru, často se tyto problémy řeší až v aplikaci. Tím pádem server odešle HTTP kód 200 a v HTML kódu je informace o tom, že stránka nebyla nalezena. A důsledkem je, že pokud budeme mít mnoho neplatných odkazů, tak vyhledávač může takovýto web navštěvovat s menší frekvencí, nebo jinak penalizovat. [15]

Všechny stránky, které chceme mít zaindexovány, musí být klasickým způsobem dostupné, odkazy tedy nesmí přijít pomocí AJAXu. Obsah, který pošleme touto metodou, je vyvolaný JavaScriptem. Musíme brát v potaz to, že vyhledávače neumí spouštět JavaScript. A pokud umí, tak pouze tak, že hledá něco, co mu připomene URL a vyzkouší ji. Když přece jenom odkazy chcete posílat později, tak je doporučováno vytvořit sitemap se všemi URL adresami. [8]

2.2.2 Technické faktory

Trocha historie: dříve se kupovali **L2 domény** podle zaměření stránek, čili podle nějakých klíčových slov. Toto se dělalo kvůli vyhledávačům, pokud našel nějaké klíčové slovo v L2 doméně, tak ji popostrčil na přední pozice. Také zvýhodňovaly národní domény, místo toho, aby zjistily, jakým jazykem jsou stránky skutečně napsány.

Naštěstí dnes jsou tyto nedostatky opraveny a programátoři vědí, že takto řazené výsledky vůbec nejsou relevantní vzhledem k hledané frázi. Proto se od těchto algoritmů naštěstí upustilo. Ovšem stále se klade velký důraz na tvar L2 domény, ovšem s jinými kritérii. Dnes se L2 doména vybírá tak, aby byla příjemná (snadno zapamatovatelná) pro uživatele nikoliv pro vyhledávače. Protože stránky a obsah se dělá pro uživatele. Pokud potřebujeme mít v L2 doméně více slov, tak nejvhodnější způsob zápisu je namísto mezery použít pomlčku (mínus, „-“). Tento znak je vyhledávači chápán jako náhrada za mezeru, tudíž název L2 domény vnitřně rozdělí.

Mnoho SEO příruček tvrdí, že je dobré mít nějaké klíčové slovo, na které chceme být ve vyhledávání vepředu. To je pravda, ale z jiného důvodu, než některé SEO příručky tvrdí. Myslí si, že vyhledávač vidí hledané slovo v L2 doméně a „hurá“ s ním do výsledků na přední pozice. Tak to ovšem není (nemělo by

být). Důvodem, proč je dobré mít v L2 doméně klíčové slovo, je to, že autoři třetích stránek často uvádějí znění L2 domény jako text odkazu. A text odkazu berou vyhledávače hodně vážně. Ovšem název L2 domény není nejklíčovější. Přibližně stejně na tom je L3 doména, kde můžete doplnit chybějící klíčové slovo. Toto se ovšem dá nahradit třeba doplněním slova do hierarchie URL.

Příklad: *bazeny.cz* – firma se rozrůstá a poptávka je hodně směřována k čerpadlům, tak upravíme a máme na výběr – *cerpadla.bazeny.cz*, nebo *bazeny.cz/cerpadla* . Vyhledávačům je jedno, kde v URL slovo nalezne. Ale všeho s mírou, nedávejte příliš mnoho klíčových slov a lomítek do URL, aby nebyla adresa příliš dlouhá a složitá. Cílem by mělo být, aby URL bylo dobře zapamatovatelné. Ovšem nezapomínejte, že větší váhu má textový obsah webové aplikace. [8], [16], [17]

Doctype je na webových stránkách jen kvůli historii z doby předchůdce HTML. Od roku 1999 slouží pouze pro přepínání vykreslovacích módů. Jiný vliv na stránky nemá. Ani v XHTML nemá jiný význam. Bohužel doctype musí být na stránce. A pokud nevíte, jaký režim chcete, tak nejjednodušší zápis bude právě to, co hledáte. Tím je nová deklarace, která vyšla s HTML5. `<!DOCTYPE html>` [18]

Hlavička dokumentu je povinný párový tag. Kromě standartních tagů (title, description, keywords), které pomáhají při SEO lze do hlavičky vložit další tagy z Open Graph protokolu. Díky několika málo tagům bude stránka zařazena do celosvětového grafu webových stránek. Také pomáhá sociálním sítím a vyhledávačům v pochopení tematického zaměření webové stránky. [19], [20]

Každá webová stránka by měla mít unikátní a správně dlouhý **popisek** (description). Ten je důležitý zejména pro vyhledávače, uživatelé tuto hodnotu standardně nevidí. Popisek by měl stručně vystihovat obsah webové stránky. Délka popisku má svá pravidla. Třeba Google délku popisku pod 40 znaků hodnotí jako krátkou a doporučuje ho rozšířit. Na druhou stranu popisek nad 155 znaků ořízne a zbytek ignoruje. Pro ostatní vyhledávače je maximální délka pro ukládání 255 znaků. [4], [8], [21]

Každá webová stránka by měla mít unikátní a správně dlouhá **klíčová slova**. Klíčová slova by měla heslovitě vystihovat obsah webové stránky. Čím jsou

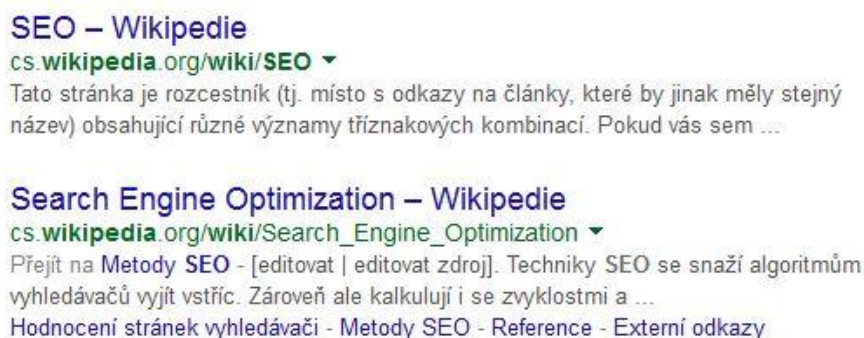
na stránce kratší (údernější) klíčová slova, tím jednotlivá slova mají větší bodové ohodnocení. Dostávají určitý počet bodů jako celek a ten se následně rozdělí mezi jednotlivá klíčová slova. Tento tag je důležitý zejména pro vyhledávače, uživatelé tuto hodnotu nevidí. Pro Google některé zdroje uvádí nedůležitost klíčových slov, jelikož obsahovaly nesmyslná slova a fráze. [4], [8]

Na začátku SEO optimalizace nebo vytváření webové aplikace je potřeba si určit všechna klíčová slova. A tyto klíčová slova následně rozdělit do co nejkonkrétnějších skupinek dle tematičnosti. Následně pro každou tuto skupinku vytvořit jednu stránku. Po této analýze a vytvoření stránek musí být vývojář schopen na jakékoliv klíčové slovo či frázi, které je vybráno, říci k jaké stránce patří, pokud toto nezvládne, tak to nemůže zvládnout ani vyhledávací algoritmus. Toto rozdělení je klíčové a ne vždy se povede na poprvé. Proto je vhodné použít nějaký nástroj pro kontrolu a sledování způsobu příchodu uživatelů do webové aplikace. Takovýmto nástrojem může být *Google Analytics*. V tomto nástroji můžeme sledovat například jaká slova či fráze uživatelé vyhledávali a díky kterým se dostaly na konkrétní stránky webové aplikace. Také zde můžeme zjistit, zda zvolená klíčová slova jsou opravdu přínosná a klíčová pro konkrétní stránky. Popřípadě se zjistí další klíčová slova, pro která může být vhodné optimalizovat obsah. Dalším zjištěním může být malá souvislost mezi jednotlivými klíčovými slovy. V tomto případě je vhodné stránku rozdělit.

Pokud autora nenapadají možná klíčová slova pro stránky a má alespoň základní klíčová slova, může použít nějaký nástroj pro klíčová slova. Tento nástroj by měl být napojen na databázi vyhledávačů, aby fráze byly aktuální. Popřípadě použít nástroj přímo od vyhledávačů jako je *Google Keyword Tool*, *Google Trends* a *Seznam statistiky hledanosti* klíčových slov. [8]

Titulek stránky je jeden z mála povinných tagů. Tento tag je důležitý pro všechny vyhledávače. Zobrazují titulek stránky ve výsledcích na prvních řádcích (Obrázek 2), mnoho uživatelů se orientuje ve výsledcích pouze pomocí titulků, někteří si i pročítají fragmenty textu ze stránky a někteří zkontrolují i URL. Do titulku stránky se umísťují nejdůležitější slova pro danou stránku webu. Titulek by měl být velice krátký, úderný a unikátní pro každou stránku. Jeho délka by měla být ideálně

od 15 do 60 znaků pro Google. Často se do titulku přidává název webu (společnosti) za pomlčkou nebo rouru (svislou čárou). [8], [21]



Obrázek 2: Výsledky vyhledávače

Metainformace pro **roboty** říká vyhledávacím robotům, jak se mají na stránce chovat. Tím je myšleno, zda může stránku indexovat a zda může pokračovat v indexaci na jiných stránkách pomocí odkazů z aktuální stránky. Tento tag nemusí být vyplněn. Pokud tak učiníte, vyhledávací roboti se chovají, jako by v tomto tagu bylo vyplněno index, follow. Tyto hodnoty jsou žádoucí a lze je kombinovat

- follow - Indexovací robot může projít odkazy na stránce
- index - Indexování stránky je povoleno
- nofollow - Indexovací robot má zakázáno procházet odkazy na stránce
- noindex - Indexování stránky je zakázáno

Nadpisy jsou nepovinné tagy. Ovšem používáním nadpisů posunete svou stránku v několika ohledech výše. Hlavním pozitivem je lepší orientace a pochopení Vaší webové stránky indexovacím robotem a screenreaderem. Obecně se doporučuje mít na stránkách alespoň nadpis úrovně 1 <h1>...</h1> a to právě jenom jednou. Měl by být celkem krátký a snažit se vystihnout stránku. Nejedná se o extrémně důležitý tag, ale určitě se navýší interní hodnocení webu, pokud se budou nadpisy používat.

Webové stránky mohou být chápány jako dokumenty. Prvním typem dokumentu může být Word nebo Writer. Tedy převážně text a nějaký grafický prvek nebo obrázek k tomu. Příkladem jiného typu dokumentu je PowerPoint nebo Impress. Mů-

že být tedy chápán jako prezentace. Mnoho obrázků, menší koncentrace textu. Ovšem ani tady by neměly chybět nadpisy a jednotná struktura pro celý web. [21]

Sémantické značky jsou celkem důležité pro pochopení informací na webových stránkách indexovacím robotem. Mnoho SEO nástrojů má test na sémantické značky, ovšem tento test je primitivně založen pouze na kontrole fontových tagů, jako jsou tučné písmo (``), kurzíva (`<i>`) a některých málo dalších. Namísto těchto značek vnucují pro tučné písmo tag `` a pro kurzívu ``. Rozdíl v definici opravdu je, ale nikde není zakázáno starší tagy používat. Tag `` slouží podle definice pouze ke grafickému zvýraznění, kdežto `` slouží i k informativnímu zdůraznění textu. Vyhledávače ovšem rozdíl mezi nimi nedělají a hodnotí tento text jako důležitější oproti ostatnímu textu.

Toto nejsou jediné sémantické značky. Pokud se pominou nové značky definované v HTML5, tak stojí za zmínku tagy `<h>` pro nadpisy, `` a `` pro seznamy. A tyto značky jsou důležitější pro pochopení informací na webových stránkách proti značkám v předešlém odstavci. Ovšem test na „potřebu“ těchto značek není tak lehké vytvořit, proto se jiné SEO nástroje nejspíš zaměřili na mnohem lehčí testování a špatnou interpretaci informací o testu a jeho důležitosti. [21]

Jsou sémantické HTML značky, které se nemusí používat a naopak značky, které by se měly používat právě pro zmíněné lepší pochopení obsahu na webové stránce. Například není nutné dávat text do odstavců (`<p>`), `<div>` je sice jen obalový prvek, avšak pro tyto účely neubírá informace o textu, prostě tam je. Kdežto pokud nebude použit seznam (``, ``) a „bude vytvořen“ pomocí odřádkování (`
`), tak tento seznam ztratí podstatnou informaci. Jako doprovodné pozitivum je lepší orientace pro zrakově postižené uživatele, kterým obsah stránek předčítá hlasová čtečka. [22]

Kontrola **uspořádání obsahu v kódu** je zbytečná z hlediska dnešních fulltextových vyhledávačů. Z minulého století se v úzké skupince HTML kodérů traduje milná informace o důležitosti pořadí obsahu a navigace. Tradovalo se, že obsah blíže k začátku je důležitější než obsah ke konci dokumentu. V dobách fulltextových vyhledávačů se tato informace nikdy nepotvrdila.

Jeden z dalších přežitých mýtů je o minimální hustotě klíčových slov v textu. Jak tomu u mýtu bývá, základní stavební zrnko je pravda. Pravdou je, že by z textu mělo být jasné, na jaká klíčová slova je text optimalizován, tudíž v nějaké formě se musí na stránce vyskytovat. Ovšem počítání hustoty je nesmyslné, nedůležité a nic neříkající.

Téměř každý zdroj informací ohledně SEO optimalizací se rozchází v radách o důležitosti jednotlivých prvků na stránkách. Ovšem převážná většina se shodne na velké **důležitosti obsahu**. Tím je míněn text na stránce, který by měl v ideálním případě být unikátní a kvalitní. Mít naprosto unikátní text není úplně možné díky velké konkurenci. Tím je myšleno, že autoři webových stránek by neměli texty mezi sebou kopírovat. Vyhledávače (obzvláště Google) si tyto informace pamatují poměrně dlouho a mohou penalizovat web se zkopírovaným obsahem. Co je to vlastně kvalitní obsah na webové stránce? Tím je myšlen právě unikátní, čtivý a zajímavý obsah. Dalším kritériem kvalitního obsahu je neopakující se text na různých stránkách webu. [8]

Jak už bylo výše zmíněno, tak optimalizace se provádí pro každou stránku webu zvlášť. Každá stránka má být optimalizována pro úzkou skupinku klíčových slov. Pokud máte jednu stránku optimalizovanou pro zimní pneumatiky, letní pneumatiky a třeba pneumatiky Continental, tak máte mnoho nepříliš souvisejících frází pro jednu stránku. Proto úzkou skupinkou je myšleno například jedna stránka pro zimní pneumatiky, další stránka pro letní pneumatiky, další pro 17 palcové pneumatiky, a tak dále. Tyto stránky mohou být vytvořeny například pomocí maticového filtru. Kde docílíme opravdu konkrétních klíčových slov pro právě jednu stránku. Příkladem může být navigace firmy prodávající pneumatiky:

- zimní pneumatiky
 - 1. výrobce pneumatik
 - pneumatiky 17 palců
 - pneumatiky 18 palců
 - 2. výrobce pneumatik
 - pneumatiky 17 palců
 - pneumatiky 18 palců

- pneumatiky 17 palců
 - 1. výrobce pneumatik
 - pneumatiky 18 palců
 - 1. výrobce pneumatik
- letní pneumatiky
 - 1. výrobce pneumatik
 - pneumatiky 17 palců
 - pneumatiky 18 palců
 - 2. výrobce pneumatik
 - pneumatiky 17 palců
 - pneumatiky 18 palců
 - pneumatiky 17 palců
 - 1. výrobce pneumatik
 - pneumatiky 18 palců
 - 1. výrobce pneumatik
- výrobce pneumatik
 - 1. výrobce pneumatik
 - zimní pneumatiky
 - pneumatiky 17 palců
 - pneumatiky 18 palců
 - letní pneumatiky
 - pneumatiky 17 palců
 - pneumatiky 18 palců
- velikost pneumatik
 - pneumatiky 17 palců
 - 1. výrobce pneumatik
 - zimní pneumatiky

Při takto navržené struktuře navigace by se mělo dávat pozor, aby nedocházelo k příliš velkému duplikování obsahu. Pokud duplikování obsahu nelze zamezit, lze říci indexovacímu robotovi, která stránka je pro web hlavní (nejdůležitější). Toho se dá docílit použitím meta tagu canonical s URL.

Textový obsah stránky by měl obsahovat širokou paletu různých i příbuzných slov. V dnešní době samozřejmě umějí vyhledávače rozpoznat různé tvary slov. Proto je velice vhodné právě ohebná slova ohýbat. A proč ohýbat slova? Je to celkem jasné. Pokud je stránka optimalizována na frázi „zimní pneumatiky“, tak šance, že uživatel vyhledává přesně tuto frázi, není stoprocentní, ani se k tomuto číslu nepřibližuje. Na druhou stranu budou často vyhledávat třeba fráze „pneu na zimu“, „pneumatiky na zimu“ a „pneumatiky pro zimní období“, zkrátka při optimalizaci lze zjistit, že uživatelé napíší do vyhledávání opravdu cokoli. Proto je skoro nutné ohýbat slova. Tím je zvýšena šance zobrazení právě vaší stránky ve výsledcích na první stránce. Některá důležitá nebo klíčová slova lze zvýraznit ztučněním nebo kurzívou. Tím se jim trošku navýší důležitost a jako se vším je potřeba zvýrazňovat se střídmostí. Jiné zvýrazňování textu jako je podtržení, přeškrtnutí, jezdící a blikající text vytváří u převážné většiny uživatelů spíše odpor než zájem ve čtení, stejně tak to vidí a chápou algoritmy vyhledávačů.

Některé SEO optimalizace k navýšení textu a počtu výskytů klíčových slov používají skrývání textového obsahu. Obsah se dá skrýt několika způsoby. První je pomocí CSS vlastnosti „display: none“, „visibility: hidden“ nebo umístit na bílou stránku bílý text. Další možností může být pomocí JavaScriptu, kde opět dosadíme CSS styly po vykreslení stránky. A nejspíše poslední způsob je překrytí textu nějakým jiným prvkem, nejčastěji netextovým. Praktiky popsané v tomto odstavci jsou zakázané. Jak moc umí Seznam a Bing tyto praktiky detekovat nikdo přesně neví, na druhou stranu se ví o Googlu, že tvrdě penalizuje weby, které používají zakázané praktiky optimalizace.

Pokud je webová aplikace vytvořená například ve Flashi, tak vyhledávač s největší pravděpodobností nemá co zaindexovat. V takovýchto případech se snaží zaindexovat alespoň něco, pak dostanou větší váhu popisek stránky, klíčová slova, title jako atribut a alternativní texty u objektů. Je-li na stránce dostatek textu, pak se zmíněným prvkům ubírá na důležitosti. [8]

Interní odkaz je odkaz jako každý jiný. Správná struktura a dostatečné množství křížových odkazů na webu zvýší hodnotu interního ranku, která může být klíčová při řazení výsledků s konkurencí. Pro rozšíření navigační struktury je velice dobrá drobečková navigace. A ideálně kontextová navigace napříč webovou aplikací, ovšem s tematickou podobností. Aby byl odkaz přínosný, měl by anchor text obsahovat slovo či slova tematicky příbuzné nebo klíčové slovo z cílové stránky. [8], [21]

Odechozí odkazy nemusejí vždy snižovat rank. Odkazováním na cizí webové stránky můžete polepšit své aplikaci, ovšem pak tento odkaz musí směřovat na silnou webovou stránku, jinak tyto odkazy mohou spíše uškodit. Samozřejmostí těchto odkazů je tematická podobnost. [8], [21]

2.3 Off-page faktory

Tyto faktory se na stránce přímo nevyskytují, tudíž autor stránky obvykle není schopen přímo tyto faktory ovlivnit, mají však přímý vliv na relevanci stránky při hledání určitého dotazu. Mezi známé off-page faktory patří zpětné odkazy, jejich množství a síla, texty těchto odkazů, případně text odkazující stránky, respektive relevance dané webové stránky vůči odkazované stránce.

Síla jednotlivých zpětných odkazů může být daná rankem odkazující stránky, počtem odkazů na této stránce, stářím odkazu, umístění na stránce, důvěryhodností stránky, důvěryhodností domény, použitím `rel="nofollow"`, typem odkazu (HTML, JavaScript atp.), různorodostí odkazujících stránek, penalizací stránky a dalšími faktory na odkazujících stránkách. [23], [24]

2.3.1 Zpětné odkazy

Ideální zpětné odkazy není tak jednoduché vytvářet. A v žádném případě uživatelé a návštěvníci webu nevědí jak správně tyto odkazy vytvářet. Jak může takový ideální zpětný odkaz vypadat? Měl by být umístěn na autoritativní nebo hodnotné stránce, tím může být například stránka výrobce nebo wikipedie. Pro Google samozřejmě platí tematická příbuznost webů, nebo alespoň stránky z které je odkazováno. Seznam naštěstí zatím tuto kontrolu neumí, ale již nějakou dobu na tom pracuje, tudíž je vhodné při vytváření zpětných odkazů myslet především na tematičnost kvůli Googlu. Dalším velkým přínosem jsou odkazy umístěné na hojně navštěvova-

ných stránkách. Nejenom kvůli předávání ranku, ale i větší pravděpodobnosti návštěvy uživatelů.

Velkým pozitivem je správný, relevantní anchor text. Zpětné odkazy typu „více informací“, „zde“ atd. jsou více méně k ničemu z hlediska anchor textu, na druhou stranu stále plní funkci odkazu a přístupu uživatelů na cílový web. Dále by bylo vhodné, aby z odkazující webové aplikace nevedlo příliš mnoho odkazů, jelikož rank aplikace se rozpočítává mezi všechny odchozí odkazy, pokud jich je hodně, může se stát, že na jeden odkaz zbyde strašně malá hodnota ranku.

Další pozitivum je jednosměrný odkaz bez `rel="nofollow"`. Díky jednosměrnému odkazu rank neteče oběma směry, nedochází k jeho úbytku a je jistota, že webová adresa nedostane penalizaci za výměnu odkazů. Penalizace samozřejmě nehrozí pro jeden obousměrný odkaz, ale může, pokud je takovýchto odkazů více. Samozřejmě ani odkaz s `nofollow` není zbytečný. Sice se nepřesune část ranku, ale pokud je odkaz na silné a hojně navštěvované stránce, tak může a nejspíše bude hodně návštěvníků chodit skrze odkaz na vaši webovou aplikaci.

Pro započnutí zpětných odkazů je ideální sociální síť a blogy či tematické diskuze. Na začátek je vhodné vytvořit stránku nebo profil webové aplikace na sociální síti. Dále přidávat nějaké články do vnitřní struktury webu. Pokud obsah v článcích a na webových stránkách bude čtivý a zajímavý, tak se nejspíš stránka (profil) začne zvětšovat díky uživatelům, kteří dále budou sami referovat a „rozšiřovat“ váš web pomocí „lajků“, „+1“, sdílení a komentářů. A právě toto je cílem. Jelikož vyhledávače přikládají stále větší váhu odkazům a signálům ze sociálních sítí jako jsou právě zmiňované „lajky“ atp. Nýbrž se jedná o „hlas lidu“.

Poslední částí ideálních zpětných odkazů je jejich cíl. Odkazy mířící na základní stránku neboli index nemají příliš velkou účinnost v porovnání s odkazy mířící někde do vnitřku webové aplikace. Je to způsobeno tím, že na indexu stránky většinou není zajímavý a konkrétní obsah. U nasbíraného ranku na index stránce můžeme doufat, že se dostatečně rozloží na ostatní stránky, které většinou vydělávají a jsou cílem optimalizace. [8], [21], [22][23], [24]

2.3.2 Ranky

Rank (hodnota) je číselná hodnota, která určuje důležitost každé stránky (každé URL samostatně). Obecně se Rank webové stránky vypočítává ze všech webových stránek, které na tuto stránku odkazují. Dále se v odkazujících stránkách určuje, zda mají podobné téma a tím se upravuje předávaná hodnota (Rank) z těchto webových stránek.

2.3.3 PageRank

Původní algoritmus PageRanku navrhl Sergey Brin a Larry Page. Právě po Pageovi byl algoritmus pojmenován PageRank (Pageovo ohodnocení). Původní algoritmus navrhli při výzkumu na Stanfordově univerzitě. A pro ověření funkčnosti svého algoritmu vytvořili i vyhledávač (později Google). Hodnota PageRanku může nabývat hodnot nula až deset.

Zjednodušeně se PageRank počítá podle toho, kolik stránek a s jakým ohodnocením na počítanou stránku odkazuje. Dále se každé převzaté ohodnocení dělí podle počtu odkazů na jiné weby. Rostoucí důležitost na internetu mají i sociální sítě. Fulltextové vyhledávače se naučili brát v potaz i signály ze sociálních sítí. Díky tomu je velice vhodné vytvořit webové aplikaci stránky na sociálních sítích, kde můžete velice snadno a rychle získat mnoho zpětných odkazů a hlavně návštěvníků.

PageRank stránky „a“ ($R_{(a)}$) se počítá přibližně takto Obrázek 3. Kde B_a je množina všech stránek odkazujících na stránku a, N_u je počet odkazů vedoucích ze stránky u. Rovnice se ještě upravuje normalizujícím faktorem c tak, aby součet všech hodnocení byl konstantní. Nelze s úplnou přesností říci, že jde o aktuální algoritmus, jelikož algoritmy si každá instituce snaží uchovat v tajemství. [25], [27], [29]

$$R(a) = c \sum_{u \in B_a} \frac{R(u)}{N_u}$$

Obrázek 3: Algoritmus pro výpočet PageRanku

2.3.4 S-Rank

S-Rank je obdoba Google PageRank. Ovšem nepoužívá stejné hodnocení. Určuje orientační hodnotu „oblíbenosti“ každé stránky na českém webu. Přesný algo-

ritmus výpočtu hodnoty je neveřejný. Hodnota S-Ranku může nabývat hodnot nula až deset.

Počítá se zejména z odkazové sítě stránek algoritmem, který zohledňuje jednak odkazy, které na stránku míří, ale i to, kam ze stránky odkazy vedou. Rostoucí důležitost na internetu mají i sociální sítě. Fulltextové vyhledávače se naučili brát v potaz i signály ze sociálních sítí. Díky tomu je velice vhodné vytvořit webové aplikace stránky na sociálních sítích, kde můžete velice snadno a rychle získat mnoho zpětných odkazů a hlavně návštěvníků.

Těmto a dalším faktorům říkáme off-page faktory. Výpočet hlavního zdroje ranku se podobá známému algoritmu Hubs & Authorities, ale je upraven tak, aby dával smysl i pro netematické skupiny stránek. [26], [28], [29], [30], [31]

2.3.5 Alexa Rank

Tento rank se podstatně liší od ostatních veřejných ranků. Tato hodnota se nepočítá z odkazů na nebo ze stránky. Alexa Rank je hodnocení domény na základě sledování návštěvnosti jednotlivých stránek uživateli, jež mají nainstalovaný speciální Alexa toolbar. Čím více se takových uživatelů podívá na Váš web, tím menší budete mít rank.

Hodnoty se pohybují v rozmezí 10,000,000 – 1 a čím je hodnota vašeho Alexa ranku vyšší, tím je postavení vašich stránek horší.

Bohužel tato hodnota je velice nepřesná, jelikož jak bylo zmíněno, musí mít uživatel nainstalovaný v prohlížeči toolbar. Je známo, že zastoupení uživatelů s tímto doplňkem není rovnoměrné po světě. A jsou i známy podvodné techniky navyšování tohoto ranku. [32], [33], [34]

2.4 Zakázané techniky

Black Hat SEO nebo také nekalé techniky SEO, např. link farmy, cloaking, doorway, duplicitní obsah, neviditelný text, a další. Snaží se většinou zvýšit podvodně rank webu a dostat se tak neoprávněně na lepší pozice ve vyhledávání. Internetové vyhledávače tyto metody považují za SPAM a takovéto stránky mohou, a tvrdě penalizují. Je několik způsobů jak vyhledávače penalizují webové stránky. Jedním z nich je snížením vnitřních hodnot v algoritmu nebo prostému posunutí ve výsledcích na-

příklad +30, +100 pozic. Tvrdší penalizací může být dočasné odstranění stránek z indexu (z výsledků vyhledávání) a v nejhorším případě trvalý BAN a tudíž odstranění z indexu, které nelze zvrátit. [35], [36], [37], [38], [39], [40], [41]

2.4.1 Skryté texty

Některé SEO optimalizace k navýšení textu a počtu výskytů klíčových slov používají skrývání textového obsahu. Obsah se dá skrýt několika způsoby. První je pomocí CSS vlastnosti „display: none“, „visibility: hidden“ nebo umístit na bílou stránku bílý text, obecně použít stejnou barvu pro text a pozadí. Další možností může být pomocí JavaScriptu, kde opět dosadíme CSS styly po vykreslení stránky a tím nějak skrýt text. Jako další způsob může být zmenšení textu na takovou velikost, že se stane nečitelným a téměř neviditelným. A nejspíše poslední způsob je překrytí textu nějakým jiným prvkem, nejčastěji netextovým. Praktiky popsané v tomto odstavci jsou zakázané. Jak moc umí Seznam a Bing tyto praktiky detekovat nikdo přesně neví, na druhou stranu se ví o Googlu, že tvrdě penalizuje weby, které používají zakázané praktiky optimalizace. [8], [42]

2.4.2 Skryté odkazy

Zakázané praktiky popsané u skrývání textu se dají použít i pro skrývání odkazů. Bohužel i toto skrývání je těžce penalizováno především vyhledávacím robotem Google. Jelikož vyhledávací robot je stroj, tak nemusí vždy přijít na skrývání odkazů nebo textu. Pokud na toto nepřijde vyhledávající robot, tak konkurence ano a záleží na ní, jak s touto informací naloží. [42]

2.4.3 Cloaking

Cloaking zpravidla vede k penalizaci od vyhledávače. Spočívá ve vytváření rozdílných stránek pro vyhledávače od těch, které jsou předkládány uživatelům. [42], [44]

2.4.4 Doorway (gateway)

Slouží k oklamání vyhledávačů. Jde o podstrčení vysoce optimalizované vstupní stránky, která je určena pouze vyhledávačům. Uživatelům v prohlížečích se nezobrazuje, neboť návštěvníky přesměrovává na jinou stránku, tohoto přesměrování si zpravidla uživatelé nevšimnou. Tato stránka s cílem zlepšit pozice ve vyhledáva-

čích obsahuje velké množství klíčových slov. Vyhledávače však tyto techniky nemají v oblibě a za jejich využití hrozí velká penalizace webových stránek. [40], [45]

2.4.5 Komentářový spam

Jedná se o psaní komentářů do blogů, příspěvků a podobných prvků za účelem vkládání nerelevantních zpětných odkazů, někdy mohou být i relevantní ale vždy jde o nežádoucí vkládání zpětných odkazů. Účelem komentářového spamu je navýšení rank hodnot, nebo alespoň větší možnost návštěvy webové aplikace pomocí prokliku. [8], [42], [43]

2.4.6 Link farmy

Link farma (odkazová farma) je speciální web s velkým množstvím stránek (v řádech stovek, tisíců, ale klidně až milionů), které jsou vzájemně propojeny odkazy. Taková webová aplikace má potom díky velkému počtu zpětných odkazů vysoký rank a zobrazuje se na prvních pozicích ve výsledcích vyhledávání. Link farmy bývají často umístěny na L3 doménách. Když je vyhledávače objeví a penalizují, majitel je jednoduše zruší. [39], [42], [46]

2.4.7 Tajné přesměrování

Tajné přesměrování využívá JavaScripty, které vyhledávací robot neindexuje, zpravidla neumí detekovat a návštěvníky přesměrovává na jiné weby. Často se jedná o stránky s nelegálním obsahem jako je warez. [8], [39], [42]

2.4.8 Automatické registrace do katalogů

Není vhodné využívat k registraci do katalogů automatizované algoritmy. Tato technika je sice těžko odhalitelná, ale je jen otázkou času, než největší vyhledávač vytvoří algoritmus na tuto detekci, jelikož vytvořené stránky jsou si dost podobné. [8], [39], [42]

2.4.9 Manipulace s rank hodnotou

Za manipulaci s rank hodnotou může být považováno například registrace do velkého počtu online katalogů. Dalším příkladem jsou PR články na free i placených PR webových aplikacích. Jelikož obě zmíněné techniky zpravidla nepřírozeň navyšují počty zpětných odkazů a tudíž i hodnotu ranku. Manipulace je tvrdě penalizována. [40], [42]

2.4.10 Microsite

Jde o webové stránky malého rozsahu, zaměřené na konkrétní produkt, produktovou řadu či službu. Jejich princip spočívá ve vybudování webových stránek s vysokou hustotou výskytu malého okruhu klíčových slov. Proto umožňují dosažení lepších pozic, než u velkého webu, který se hůře optimalizuje. Uživatelé se po vyhledání dotazu nejprve dostanou na microsite, která je následně může sama přesměrovat, nebo každý odkaz míří na hlavní webovou aplikaci. Tyto stránky mají lepší pozice ve vyhledávačích a poskytují kvalitní a relevantní zpětné odkazy. Ale bohužel vyhledávací robot penalizuje jak microsites, tak i hlavní webovou aplikaci, což je velký problém. [8], [47]

2.5 Nette Framework

V oblasti vývoje webových aplikací existuje velké množství knihoven obsahující zobecněné moduly, funkce a třídy, souhrnně pojmenované jako framework.

Nette Framework (dále jen Nette) je open source nástroj pro vytváření webových aplikací v PHP jazyce verze 5. Nette je svobodný software, nabízený pod licenci *New BSD*, nebo *GNU GPL* verze 2 nebo 3, licenci si může uživatel vybrat sám a následně ctít. Tento framework začal vyvíjet jeden český programátor a zakladatel David Grudl. Nyní se na vývoji podílí stále zakladatel a Nette Foundation, díky fóru a diskuzím se na vývoji, ošetřování chyb a vytváření doplňků podílí i uživatelé. V současné době se jedná o jeden z nejpoužívanějších PHP frameworků nejen v České republice. [48], [49]

2.5.1 Výhody

Nette využívá událostmi řízené programování a celý se skládá z malých poměrně nezávislých komponent, které se ovšem dají použít nezávisle při vytváření vlastní aplikace. Mezi hlavní výhody patří eliminace výskytu bezpečnostních děr a jejich zneužití, jako je např. XSS, CSRF, session útoky atd. Díky neustálému vývoji tento moderní framework vlastní podporu AJAX, AJAX, architekturu pro vytváření MVC aplikací, propracovanou strukturu pro routování uživatelsky přívětivých URL adres oběma směry, tudíž odpadá nutnost modifikace a vytváření .htaccess souborů. Díky cachování výsledků a dalších částí nejenom při sestavování latte (vzhledu) se jedná o jeden z nejrychlejších frameworků. Snadná je i rozšiřitelnost díky doplňkům,

kteře jsou přímo na oficiálních stránkách. Jednou důležitou komponentou Nette jsou ladící nástroje, díky kterým odstraňování chyb je radostí. [48], [50]

2.5.2 Požadavky

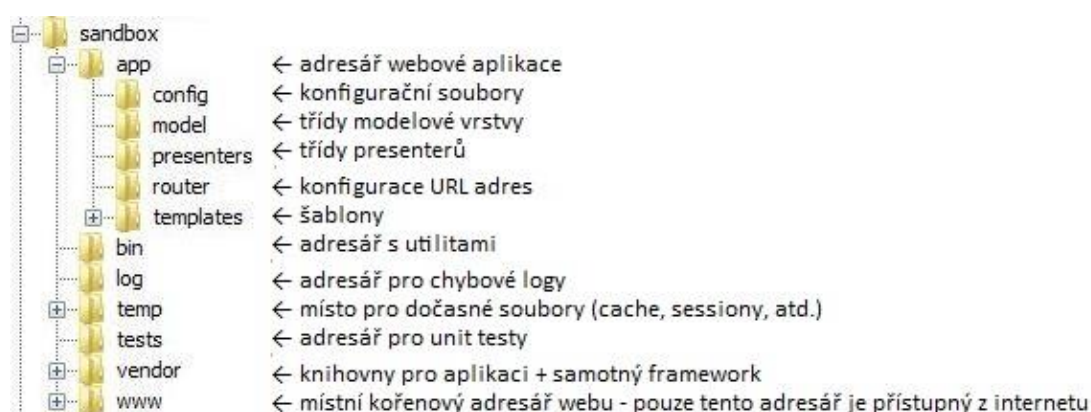
Nette Framework vyžaduje pro užívání minimálně PHP verzi 5.3.1 a vyžaduje některé moduly a nastavení (správné a bezpečné) webového serveru. Hlavními požadavky jsou verze PHP jazyka, `mod_rewrite`, `ini_set()`, Multibyte String extensi-on a zpětná kompatibilita s PHP 4 musí být vypnutá. Pro kontrolu požadavků je k dispozici nástroj umístěný v každé Nette distribuci ve složce *tools/Requirements-Checker*, který stačí pouze spustit na cílovém webovém serveru. [51]

2.5.3 Stažení a instalace

Existují dva způsoby jak Nette stáhnout a nainstalovat. Prvním z nich je pomocí *Composeru*, kterému předáme název projektu a cíl. On se následně postará o stažení a správné umístění. Samozřejměstí *Composeru* je i stahování jednotlivých modulů. Jako další způsob je ruční stažení archivu z oficiální stránky a nakopírování souborů do složky. [52]

2.5.4 První spuštění

Po nainstalování Nette na webový server lze spustit tzv. *Sandbox*. Jde o předpřipravenou úvodní stránku s velkým textem *Congratulations!*, která oznámí správné nakopírování a fungování frameworku.



Obrázek 4: Struktura adresářů Sandboxu

Do místního kořenového adresáře webu se ukládají CSS styly, JavaScripty, obrázky a další veřejně přístupné soubory. Díky umístění všech veřejně přístupných souborů v jedné složce lze nastavit kořenový adresář aplikace do této složky.

Však nejdůležitější složkou a složkou ve které se tráví nejvíce času je *app/*. Ta obsahuje konfigurační soubor pro moduly, továrničky a přístupy do databáze v souboru *config.neon*, také nastavení routovacích pravidel a vlastní naprogramované presentery a Latte šablony. Dále v této složce je jeden soubor a tím je *bootstrap.php*, který slouží k zavedení automatického načítání tříd pomocí *namespace*. Díky tomu se nemusí ručně řešit načítání tříd. Také načítá celý Nette Framework, nastavuje se zde produkční a vývojový režim aplikace, adresář s umístěním logů, dočasných souborů jako je cache atd. [52]

2.5.5 MVC aplikace

Při správném používání Nette je nedílnou součástí využívat softwarovou architekturu typu MVC. Jak bylo zmíněno, tak důležité soubory aplikace jsou k nalezení v adresáři *app/*, kde jsou umístěny další adresáře pro každou část softwarové architektury. Modelová část aplikace je umístěna v adresáři *model/*, vzhled je uložen v adresáři *templates/* a poslední část se umísťuje do adresáře *presenters/*. Díky této architektuře se snadno dělají testy jednotlivých částí zvlášť.

Modul je pouze způsob členění zdrojových kódů. Pokud je vytvářena složitější aplikace, tak jednotlivé pohledy a kontrolery lze přesunout do různých modulů, které lze svévolně vytvářet. Často se moduly používají pro odlišné části webové aplikace. Jeden modul může obsluhovat třeba část pro uživatele a další administrativní část. Moduly obsahují své vlastní presentery a pohledy, které mohou mít shodné názvy napříč aplikací.



Obrázek 5: Použití modulů v aplikaci

Díky modulárnosti může být modelová část buď centralizována (viz Obrázek 5). Při velice odlišných modelech pro jednotlivé moduly lze model také přesunout do každého modulu. [55]

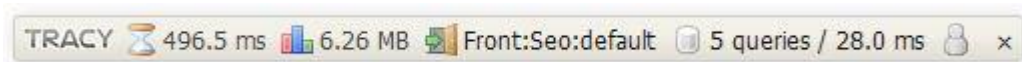
2.5.6 Ladící nástroj Tracy

Velice důležitým modulem v Nette je ladící nástroj *Tracy*. Z minulosti známý jako *Nette\Diagnostics\Debugger*. Během své historie získala název *Laděnka* [53]. Pomáhá programátorovi v

- Rychlému odhalení chyb (jak ve vývojovém, tak v produkčním režimu)
- Logování chyb v produkčním režimu
- Přehledném a formátovaném výpisu proměnných
- Měření časů v aplikaci

Tento nástroj je pro vývoj PHP aplikací velice vhodný, jelikož PHP jazyk je jazykem se slabou typovou kontrolou a obecně řečeno jde o volný jazyk. Proto se při vývoji vyskytují velice často chyby, které není tak jednoduché přesně identifikovat ze základního chybového výpisu.

Další užitečnou součástí *Tracy* je ladící panel neboli *Debugger Bar* (Obrázek 6). Tento panel se zobrazuje v pravém dolním rohu každé stránky, pokud je aplikace nastavena ve vývojovém režimu.



Obrázek 6: Tracy panel

Panel obsahuje několik důležitých informací. První je celkový čas aplikace, který byl potřeba pro kompletní sestavení, další položkou je využitá paměť během sestavování. Dále je vidět v jakém presenteru a pohledu se právě aplikace nachází, při rozkliknutí této položky jsou vidět všechny routovací pravidla a povolené permutace, také zde jsou k nalezení případné parametry předané v URL. Předposlední částí jsou informace o počtu dotazů a jejich celkový čas na straně databáze, tato část obsahuje k nahlédnutí všechny dotazy použité aplikací a jejich odpovědi s jednotlivými časy. Poslední položkou panelu je identita *Nette\Security\Identity*, pokud je vývojář

aktuálně přihlášen, jsou zde k nahlédnutí informace ohledně identity. Jednotlivé položky se zobrazují v závislosti na konkrétním nastavení aplikace. Výchozí částí identity je id uživatele, jeho role (práva), email, jméno a příjmení (Obrázek 7). Pokud je aplikace nastavena v produkčním režimu, tak se panel nezobrazuje a případné chyby jsou logovány do předem nastaveného textového souboru. Samozřejmostí panelu je jeho modulárnost pomocí doplňků, jak existujících, tak i vlastních.



Obrázek 7: Nette identita

2.5.7 Šablonovací jazyk Latte

Jako další nesmírně užitečná část Nette je intuitivní šablonovací systém *Latte*. Důvodem vytváření šablonovacích systémů v PHP je nepřehledný kód při vytváření vzhledů, který neustále obsahuje úvodní a koncové značky PHP `<?php ... ?>`. A případně i funkce pro převod speciálních znaků na ASCII kód.

Latte disponuje velmi důležitou vlastností v podobě výpisu proměnných, které jsou automaticky escapovány. Výpis libovolné proměnné by měl být vždy proveden pomocí escapování, aby bylo dodrženo převedení znaků obsahující v HTML speciální význam na alternativní zápis. Bez escapování proměnných by mohla v aplikaci vzniknout bezpečnostní díra v podobě XSS. Další užitečnou vlastností jsou helpery a makra, které dokáží zpříjemnit a usnadnit práci díky možnosti jejich vytváření a používání. [54]

2.5.8 Databáze

Tento framework obsahuje databázovou vrstvu pro zjednodušení práce s databázemi, podporuje širokou paletu relačních databází (Obrázek 8). Základní třída

pro připojení *Nette\Database\Connection* obaluje PDO rozhraní a slouží k připojení k databázi. Pokud programátor preferuje vlastní psaní a skládání SQL dotazů, pak může využít *Nette\Database\Context*. A vrcholem těchto databázových vrstev je *Nette\Database\Table*, které optimalizuje a zjednodušuje výběr dat. Pro správné fungování databázové vrstvy je třeba mít nastavení databázového uložiště jako InnoDB, které podporuje cizí klíče. [55], [56]

Databázový server	DSN jméno	Podpora v Database	Podpora v Database\Table
MySQL	mysql	ANO	ANO
PostgreSQL	pgsql	ANO	ANO
Sqlite 3	sqlite	ANO	ANO
Sqlite 2	sqlite2	ANO	-
Oracle	oci	ANO	-
MS SQL (PDO_SQLSRV)	sqlsrv	ANO	ANO
MS SQL (PDO_DBLIB)	mssql	ANO	-
ODBC	odbc	ANO	-

Obrázek 8: Nette - podporované databáze

2.6 Twitter Bootstrap

Twitter Bootstrap (dále jen Bootstrap) je hojně používaný a velice oblíbený framework pro stylování moderních webových aplikací. Bootstrap je vydáván pod licencí *MIT* [57]. [58]

Nette Framework obsahuje už v samotném jádře podporu pro Bootstrap v rendrovacích metodách. Tyto metody se dají upravit pro konkrétní využití v různých aplikačních příležitostech. Díky této podpoře lze rendrovací metody upravit, přetížít a tím nastavit jednotný vzhled pro celou aplikaci. Samozřejmostí je úprava vzhledu při vytváření konkrétních formulářů.

2.7 MySQL

MySQL je relační databázový systém pod dvojí licencí, jednak bezplatná *GPL* [59] a komerční placená. Právě díky bezplatnému použití se jedná o nejrozšířenější databázový systém, který je navíc multiplatformní. Velice oblíbená kombinace

na webových serverech je LAMP (Linux, Apache, MySQL, PHP), popřípadě WAMP (Windows, Apache, MySQL, PHP). [60]

2.7.1 MyISAM

Protože MyISAM je jeden z nejstarších formátů úložiště dat, které byly zařazeny do MySQL, obsahuje velké množství funkcí, které byly vyvinuty během let jeho existence. Také se jedná o velice rychlé úložiště. Velká nevýhoda tohoto úložiště je nepodporování cizích klíčů jako integritního omezení a transakcí. [61]

2.7.2 InnoDB

InnoDB je jeden z několika dalších formátů úložiště dat, které byly zařazeny do MySQL. Mezi klíčové vlastnosti tohoto úložiště je podpora transakcí a cizích klíčů jako integritního omezení. [62]

3 Cíl práce

Cílem bakalářské práce je stanovení důležitých a významných faktorů z hlediska SEO optimalizace webových stránek pro nejpoužívanější české vyhledávací roboty. Tyto faktory se od sebe moc neodlišují, obecně se dá říci, že povolené praktiky a postupy pro získání vyšších pozic ve vyhledávacích službách na internetu jsou stejné nebo velice podobné.

Tyto povolené praktiky a postupy je potřeba sjednotit a vytvořit z nich univerzální řešení pro analýzu zdrojového kódu na webových stránkách. Na základě takového postupu vytvořit samostatnou webovou aplikaci, která dokáže tento postup implementovat a analyzovat jakoukoli internetovou stránku.

Tento webový nástroj nebude v žádném případě používat bodového nebo procentuálního ohodnocení pro interakci správnosti SEO optimalizace, také nebude využívat násilného vnucování výsledků a informací.

Pro vytvoření webové aplikace byl vybrán Nette Framework pro jeho snadnou implementaci, vytvoření struktury, bezpečnost, práci s databází a formátovanému výstupu chyb při vytváření aplikace.

Aplikace bude obsahovat testy nejen pro praktiky SEO optimalizace, ale i pro kontrolu postupu ve výsledcích. Dále bude sjednocovat některé nástroje vyhledávačů. Testy budou rozděleny do tematických částí. Tím se zajistí menší náročnost na server, uživatelé nebudou muset čekat tak dlouho na výsledky a nedojde k tak rychlému dosažení limitů u vyhledávačů.

Webovou aplikaci bude možné využívat na libovolném zobrazovacím zařízení, od monitorů, přes tablety až po mobilní telefony. Toto bude zajišťovat moderní framework pro vytváření responzivních vzhledů.

4 Postup vývoje aplikace

Po nastudování různých SEO optimalizačních technologií jako jsou povolené a zakázané SEO praktiky, On-page a Off-page faktory. Byl vytvořen prvotní návrh testů, tento návrh byl během vývoje několikrát rozšířen. Tudiž se dá říct, že návrh a realizace řešení probíhaly poměrně paralelně.

Aplikace kromě frameworků (Nette a Bootstrap) využívá pouze nutné API pro komunikaci s některými vyhledávači, jelikož nepodporují klasické dotazy pomocí jiných přístupů. Cílem je vytvořit vlastní řešení a postupy pro testy, kde nebudou použity cizí softwarové technologie.

4.1 Použité technologie

Aplikace je psaná v programovacím jazyce PHP s databázovým systémem MySQL. Tyto technologie byly použity díky své obrovské rozšířenosti a podpory na drtivé většině webhostingů.

Dále jsou použity dva frameworky. Prvním je Nette Framework díky svému zabezpečení, výkonu a aktivní velké komunitě, a díky čemu lze nalézt mnoho vyřešených problémů. Druhý framework je Twitter Bootstrap, který slouží k rychlému vytváření responzivních vzhledů webových aplikací pomocí CSS stylů a rozšířených grafických funkcí pomocí JavaScriptu. V této kombinaci lze celkem snadno vytvořit moderní vzhledy.

4.2 Kontrola minimálních požadavků

Před samotným vývojem aplikace je nutné zajistit, aby byly splněny alespoň minimální požadavky pro běh Nette Frameworku. Jelikož je celá aplikace vytvářena pomocí Nette, je nutné, aby server splňoval minimální požadavky pro jeho spuštění. Součástí distribučního balíčku Nette je samostatný nástroj nazvaný *Requirement-Checker* [51]. Tento nástroj po nahrání na server otestuje nainstalované moduly, konfiguraci serveru a informuje o možnostech a omezeních využití funkcí frameworku. Pokud nástroj zobrazí zelený panel s textem *Congratulations! Server configuration meets the minimum requirements for Nette Framework.* [51], tak lze Nette spustit. Ovšem stále některé moduly nemusí fungovat, jelikož nejsou důležité pro samotný běh Nette, ale rozšiřují jeho možnosti.

4.3 Návrh databáze

Na základě zjištěných informací o SEO optimalizačních technikách a částí, na které je potřeba se zaměřit byl vytvořen návrh databázové vrstvy, který byl v průběhu vývoje několikrát rozšířen dle aktuálních potřeb a rozšiřování aplikace.

Databáze obsahuje třináct tabulek (viz. Obrázek 9):

- uzivatele
- adresy
- datumy
- zdrojovy_kod
- ranky
- pozice_vyraz
- zpetne_odkazy
- pocet_zaindexovanych
- kategorie_tematu
- podkategorie_tematu
- zpetne_odkazy_diskuze
- zpetne_odkazy_chat
- nastaveni

První tabulka s názvem uzivatele obsahuje informace o registrovaných uživateli. Výchozím uživatelem pro testy v neregistrované části je první uživatel, který je fiktivním a je v databázi pouze z nutnosti integritních omezení. V tabulce uzivatele jsou sloupce email, heslo, jmeno, prijmeni a role.

Tabulka adresy obsahuje sloupce uzivatele_id a adresa. Atribut uzivatele_id je cizím klíčem z tabulky uzivatele, druhý atribut adresa slouží k ukládání URL adres, které jsou uživatelem zadány k testování.

Další tabulkou je datumy, tato tabulka rozšiřuje adresy o datum provedení testu.

Obsahem tabulky zdrojovy_kod jsou vazební atributy k tabulkám adresy a datumy, dalším atributem je kod, do kterého se ukládá celý zdrojový kód testované webové stránky. Posledním sloupcem je čas obsahující čas načtení zdrojového kódu.

Tabulka ranky opět obsahuje vazební atributy k tabulkám adresy a datumy. Dále jsou součástí další dva atributy. A to google obsahující PageRank hodnotu a seznam obsahující S-Rank hodnotu. Není zde AlexaRank, jelikož při zjišťování hodnoty lze zjistit postup tohoto ranku.

Obsahem tabulky pozice_vyraz je ukládání první pozice ve vyhledávacích na uživatelem zadané slovo či frázi. Opět jsou součástí vazební atributy na tabulky adresy a datumy. Do atributů google, seznam a bing se ukládají pozice hledané stránky. Další skupinkou atributů je googleCelkem, seznamCelkem a bingCelkem, kam se ukládá celkový počet nalezených výsledků z daného vyhledávače.

Tabulka zpetne_odkazy obsahuje vazební atributy na tabulky adresy a datumy. Dále sloupečky google, seznam a bing kam se ukládají počty zpětných odkazů podle daného vyhledávače.

Obsahem tabulky pocet_zaindexovanych jsou opět vazební atributy na tabulky adresy a datumy. Dalšími atributy jsou google, seznam a bing, kam se ukládají počty zaindexovaných stránek v jednotlivých vyhledávacích.

Další sadou tabulek jsou tabulky k diskuzi, která je zaměřená ke komunikaci mezi registrovanými uživateli za účelem rozumné výměny tematických odkazů. První tabulkou je kategorie_tematu, která obsahuje atribut nazev, do kterého se ukládají tematické okruhy.

Tabulka podkategorie_tematu je závislá na předešlé tabulce a zužuje tematické okruhy pro lepší tematickou příbuznost.

Pro ukládání diskuzí slouží tabulka zpetne_odkazy_diskuze. Tato tabulka je závislá na podkategorie_tematu, kategorie_tematu a adresy. Dále obsahuje název diskuze v atributu nazev a datum s časem vytvoření v atributu datum_cas.

Do tabulky zpetne_odkazy_chat se ukládají jednotlivé chatovací zprávy do sloupečku text a do datum_cas se ukládá datum s časem vytvoření zprávy. Tato tabulka má vazby na adresy, podkategorie_tematu a kategorie_tematu.

The diagram illustrates the database schema for 'seo_pomoc'. It features 11 tables, each with its fields and data types listed. The tables are interconnected by lines representing relationships:

- seo_pomoc.uzivatele** (id: mediumint(8) unsigned, email: varchar(250), heslo: char(60), jmeno: varchar(60), prijmeni: varchar(60), role: tinyint(3) unsigned) is linked to **seo_pomoc.adresy** and **seo_pomoc.pozice_vyraz**.
- seo_pomoc.adresy** (id: mediumint(8) unsigned, uzivatele_id: mediumint(8) unsigned, adresa: varchar(255)) is linked to **seo_pomoc.zdrojovy_kod**, **seo_pomoc.pozice_vyraz**, **seo_pomoc.pocet_zaindexovanych**, **seo_pomoc.zpetne_odkazy**, **seo_pomoc.nastaveni**, **seo_pomoc.ranky**, **seo_pomoc.zpetne_odkazy_diskuze**, and **seo_pomoc.zpetne_odkazy_chat**.
- seo_pomoc.zdrojovy_kod** (id: int(10) unsigned, adresa_id: mediumint(8) unsigned, datum_id: int(10) unsigned, kod: mediumtext, cas: smallint(5) unsigned) is linked to **seo_pomoc.pozice_vyraz** and **seo_pomoc.pocet_zaindexovanych**.
- seo_pomoc.pozice_vyraz** (id: int(10) unsigned, adresa_id: mediumint(8) unsigned, datum_id: int(10) unsigned, vyraz: varchar(255), google: tinyint(4) unsigned, seznam: tinyint(4) unsigned, bing: tinyint(4) unsigned, googleCeklem: int(10) unsigned, seznamCeklem: int(10) unsigned, bingCeklem: int(10) unsigned) is linked to **seo_pomoc.pocet_zaindexovanych** and **seo_pomoc.zpetne_odkazy**.
- seo_pomoc.datumy** (id: int(10) unsigned, datum: timestamp, adresa_id: mediumint(8) unsigned) is linked to **seo_pomoc.pocet_zaindexovanych** and **seo_pomoc.zpetne_odkazy**.
- seo_pomoc.pocet_zaindexovanych** (id: int(10) unsigned, adresa_id: mediumint(8) unsigned, datum_id: int(10) unsigned, google: mediumint(8) unsigned, seznam: mediumint(8) unsigned, bing: mediumint(8) unsigned) is linked to **seo_pomoc.zpetne_odkazy** and **seo_pomoc.nastaveni**.
- seo_pomoc.zpetne_odkazy** (id: int(10) unsigned, adresa_id: mediumint(8) unsigned, datum_id: int(10) unsigned, google: mediumint(8) unsigned, seznam: mediumint(8) unsigned, bing: mediumint(8) unsigned) is linked to **seo_pomoc.nastaveni** and **seo_pomoc.ranky**.
- seo_pomoc.nastaveni** (id: mediumint(8) unsigned, adresa_id: mediumint(8) unsigned, kategorie_tematu_id: tinyint(3) unsigned, podkategorie_tematu_id: smallint(5) unsigned, vymena_odkazu: bit(1)) is linked to **seo_pomoc.ranky** and **seo_pomoc.zpetne_odkazy_diskuze**.
- seo_pomoc.ranky** (id: int(10) unsigned, adresa_id: mediumint(8) unsigned, datum_id: int(10) unsigned, google: tinyint(4) unsigned, seznam: tinyint(4) unsigned) is linked to **seo_pomoc.zpetne_odkazy_diskuze** and **seo_pomoc.zpetne_odkazy_chat**.
- seo_pomoc.zpetne_odkazy_diskuze** (id: int(10) unsigned, adresy_id: mediumint(8) unsigned, podkategorie_tematu_id: smallint(5) unsigned, kategorie_tematu_id: tinyint(3) unsigned, nazev: varchar(255), datum_cas: datetime) is linked to **seo_pomoc.zpetne_odkazy_chat** and **seo_pomoc.podkategorie_tematu**.
- seo_pomoc.zpetne_odkazy_chat** (id: int(10) unsigned, zpetne_odkazy_diskuze_id: int(10) unsigned, adresy_id: mediumint(8) unsigned, text: text, datum_cas: datetime) is linked to **seo_pomoc.podkategorie_tematu**.
- seo_pomoc.podkategorie_tematu** (id: smallint(5) unsigned, kategorie_tematu_id: tinyint(3) unsigned, nazev: varchar(255)) is linked to **seo_pomoc.kategorie_tematu**.
- seo_pomoc.kategorie_tematu** (id: tinyint(3) unsigned, nazev: varchar(255)) is the final table in the hierarchy.

4.4 Postup vytváření testů

Testy faktorů umístěných přímo na stránce byly programovány přibližně tak, jak se vyskytují na stránkách. Tudíž začínalo se hlavičkou dokumentu a dále se pokračovalo vytvářením testů na obsah stránky.

4.5 On-page faktory

Jak již bylo zmíněno, tyto faktory se vyskytují přímo na webové stránce a proto je programátor schopen tyto dílčí prvky upravit a tím zlepšit hodnocení webové stránky.

On-page faktory se testují na jedné stránce výsledné aplikace, tím je zajištěn menší přenos dat a menší počet dotazů a spojení na cizí webové servery. Sjednocení testů na jednu stránku je důvodem právě zmíněného počtu dotazů a odpovědí. Pokud je aplikace umístěna na webhostingu, tak se může stát, že při velkém a rychlém počtu odpovědí se s největší pravděpodobností lze setkat nejdříve s upozorněním od majitele na DoS útok směřovaný na Vaši adresu. A v horším případě může zablokovat danou aplikaci.

4.5.1 Získání zdrojového kódu

Pro načtení zdrojového kódu byla zvolena jednoduchá funkce, která načítá pouze obsah stránky a vloží ho do proměnné (viz. Obrázek 10). Tuto funkci lze rozšířit o informace o prohlížeči. Díky čemuž se dotaz tváří jako by byl odeslán prohlížečem (uživatel), nikoliv programem, jelikož některé webové aplikace vracely prázdný dokument a po přidání informací o prohlížeči bylo vše v pořádku. Toto bylo zjištěno až v průběhu vývoje.

```
private function ziskejZdrojovyKod($adresa, $pocitac) {
    $opts = array(
        'http'=>array(
            'method'=>"GET",
            'header'=>"Accept-language: en\r\n" .
                "Cookie: foo=bar\r\n",
            'user_agent'=>$_SERVER['HTTP_USER_AGENT']
        )
    );

    $context = stream_context_create($opts);
    $zdrojovyKod = @file_get_contents("http://".$adresa, false, $context);
    if($zdrojovyKod != false)
        return " " . $zdrojovyKod;
    return false;
}
```

Obrázek 10: Načtení zdrojového kódu

4.5.2 Doctype

Pro tento test bylo potřeba zjistit obecný tvar doctype, který se objevuje ve všech deklaracích. Tímto tvarem je po převedení textu na malá písmena „<!doctype ...>“. A následné převedení na originální velikost textu.

4.5.3 Hlavička

I když se na hlavičce jako celku neprovádí žádné testy, je zařazena do výpisu. Hlavním důvodem vypisování hlavičky je kontrola uživatelem, zda algoritmus našel hlavičku a dokázal správně určit začátek a konec, jelikož bez tohoto určení některé následující testy nemohou být úspěšně provedeny. Dalším důvodem je připojení rad o prvcích, které je vhodné do hlavičky nezapisovat a možnosti rozšíření o meta tagy Open Graph. Tyto tagy mohou vyhledávači říci doplňující informace o stránce.

4.5.4 Kódování stránky

Kódování webové stránky není důležité z ohledu SEO optimalizace. Na druhou stranu je důležitý pro všechno ostatní. Pokud kódování není vyplněno, tak vyhledávač neví jak s daným textem zacházet a při práci s ním může text obsahovat nesmyslné znaky a tím pádem i slova a stránka může získat penalizaci za nesmyslný obsah. Dále je důležitý pro tuto aplikaci, aby mohla správně pracovat s textem na stránce. Jelikož nejčastějším kódováním je UTF-8, tak při absenci této informace na stránce je automaticky počítáno s tímto kódováním.

Aplikace je kompletně napsána v kódování UTF-8 jelikož to vyžaduje Nette. Stránky využívající jiné kódování jsou překódovány právě do tohoto kódování, aby aplikace byla schopna s textem plnohodnotně pracovat.

Jelikož metainformace o kódování stránky prošla změnou v HTML5, tak algoritmus hledá buď starý zápis (dlouhý), nebo nový, který je podstatně zkrácen.

4.5.5 Titulek

Titulek, jak už bylo popsáno v teoretické části, je jeden z nejdůležitějších faktorů z hlediska SEO. Jelikož titulek stránky je párový tag, tak jeho nalezení není obtížné, ovšem pokud jsou v hlavičce vloženy JavaScripty, tak jsou nejdříve odstraněny a to z důvodu, aby se vypsál titulek, který není dosazen scriptem. Algoritmus vyhledává počáteční a koncovou značku tohoto tagu a následně získá jejich obsah, který je

vypsán uživateli. Také se počítá jeho délka pro vytvoření upozornění na jeho případně nevyhovující délku a to jak krátkou, tak dlouhou.

4.5.6 Metainformace

Dalším výpisem jsou metainformace. Toto slouží pro uvědomění programátora, zda hlavička obsahuje všechny meta tagy, které je vhodné na stránce mít.

Zde je snaha, aby si tyto nedostatky uvědomil uživatel sám bez upozornění, jelikož vědomosti nabitě sebou samým jsou daleko cennější, než „naservírované“ informace. Dalším důvodem je, že tyto značky jsou separovány od zbytku kódu psaného uživatelem, tudíž sám snáze nalezne nedostatky.

4.5.7 Popisek

Popisek (description), jak už bylo popsáno v teoretické části, není z hlediska hodnocení stránky vyhledávači tak důležitý, ale nese určitou váhu důležitosti a vzhledem k tomu, že je občas používán jako popisek stránky ve výsledku vyhledávání, byl zahrnut do analýzy. Byla k tomuto prvku přidána rada k jeho délce, která by měla být v ideálním případě mezi 40 a 155 znaky pro Google a do 255 znaků pro ostatní fulltextové vyhledávače.

4.5.8 Klíčová slova

Test na klíčová slova (keywords) pouze ověřuje, zda tento meta tag existuje. Jelikož jeho důležitost je velice sporná a je známo, že Google tuto metainformaci naprosto ignoruje. Na druhou stranu uvádění klíčových slov nesnižuje ohodnocení a může popřípadě pomoci. Tudíž je vhodné je uvádět, jelikož programátor vidí, na která slova či téma má být stránka zaměřená. Zde jsou uváděny pouze obecné informace o údernosti, tematičnosti a krátkosti.

4.5.9 Autor

Tento prvek je přidán do výpisu pouze z důvodu zviditelnění se, který může pomoci napříč webovými aplikacemi, pokud vytváříte kvalitní webové stránky s kvalitním obsahem.

4.5.10 Informace pro roboty

Metainformace „robots“ není důležitá pro SEO, ale pokud je vyplněna špatně, tak může nechtěně zabránit robotům v indexaci a průchodu webovou stránkou. Je zde vytvořena rada jak může vypadat tento tag a jakých hodnot může nabývat. Také je zmíněno, že absence této informace je chápána jako povolení indexace a sledování odkazů mířících z této konkrétní stránky.

4.5.11 Nadpisy

U nadpisů v dnešní době není moc co testovat. Měly by být na stránce, jelikož určitým způsobem pomáhají v ohodnocení, ale nejedná se o klíčové prvky. Také není potřeba dodržovat sémantickou hierarchii. Test ukazuje, zda nadpis obsahuje klíčová slova a která. Pokud nadpis klíčová slova neobsahuje, tak se zas tolik neděje, ale měl by charakterizovat text, který ho následuje. Dále je ve výpisu vidět jaké texty jednotlivé nadpisy obsahují a je vypsán i kompletní obsah nadpisů pro uvědomění čistoty kódu.

4.5.12 Odkazy

Odkazy jsou stále nejdůležitějším prvkem pro dobré umístění ve výsledcích. Jak bylo zmíněno v teoretické části zaměřené na On-page faktory. I vnitřní odkazy velice pomáhají a jsou užitečné. Také mohou pomoci odkazy mířící z dané stránky na cizí webové aplikace. Aplikace zobrazuje všechny odkazy na stránce s jejich umístěním podle čísla řádku ve zdrojovém kódu.

Kvůli stále zlepšujícím algoritmům na detekci odkazových panelů a podobných zakázaných praktik s odkazy je přidáno zvýrazňování odkazů mířících mimo webovou aplikaci. Prvním je barevné zvýraznění pro odkazy obsahující `rel="nofollow"`. Další zvýraznění je při absenci tohoto atributu, a pokud je takovýchto odkazů v řadě více jak tři, tak jsou označeny jako možný odkazový panel, který je velmi penalizován vyhledávači.

Dále je přidán počet všech odkazů na stránce a počet interních a externích odkazů.

4.5.13 Návrh klíčových slov

Tento test počítá výskyty slov na stránce. Jelikož není volně dostupný žádný algoritmus na převod ohebných slov v českém jazyce na základní tvary (1. pády, infinitivy), tak se sčítají pouze naprosto shodná slova. Dále jsou ke každému slovu připojeny počty výskytů a odkazy na služby Seznam a Google, kde vyhledávače mají volně přístupné statistiky hledanosti slov a frází, díky čemu se uživatel může rychle a pohodlně podívat jak jsou daná slova vyhledávána uživateli těchto vyhledávačů. Bing bohužel tyto služby nenabízí, proto nejsou připojeny odkazy na tento vyhledávač. Dále algoritmus zjišťuje, zda jednotlivá slova jsou v klíčových slovech a při shodě toto slovo zvýrazní.

4.5.14 Kontrola klíčových slov v URL

Pokud vyhledávač nalezne klíčové slovo v URL, tak přidává nějaké bodíky dané stránce, proto je i tento test přidán. Výsledkem tohoto testu je informace, zda je či není nějaké klíčové obsažené v URL adrese.

4.5.15 Alternativní texty u obrázků

Aplikace vypisuje všechny alternativní texty u obrázků a číslo řádku ve zdrojovém kódu. Pokud není vyplněn tento atribut, tak je vypsán text s upozorněním na absenci této informace.

4.5.16 Délky

Tyto informace slouží pro kontrolu a uvědomění si jak moc textu je stránce. Jak bylo zmíněno, množství a kvalita textu je velice důležitá. Také je vhodné ve zdrojovém kódu mít pokud možno pouze značkovací jazyk s textem. Právě kvůli těmto důvodům jsou přidány tyto výpisy:

- velikost zdrojového kódu
- velikost JavaScriptu v HTML kódu
- velikost CSS v HTML kódu
- celkový počet slov na stránce
- počet textových znaků na stránce

4.5.17 Čas načtení zdrojového kódu

Jelikož je známo, že uživatelé stejně tak jako Google preferují stránky s rychlou odezvou, tak je přidána i tato položka. Zde se vypíše čas potřebný pro vygenerování a přenos zdrojového kódu. Ten je barevně rozdělen do pěti kategorií:

- do 1,5 sekundy
- od 1,5 do 3 sekund
- od 3 do 5 sekund
- od 5 do 8 sekund
- nad 8 sekund

4.5.18 robots.txt

Tento test zjišťuje existenci tohoto souboru a informuje o zjištění uživatele. Soubor není klíčový z hlediska SEO, ale při nesprávném nastavení může blokovat některé stránky webu, které blokovány být nemají.

4.5.19 sitemap.xml

Tento test zjišťuje existenci tohoto souboru a informuje o zjištění uživatele. Existence tohoto souboru pomáhá indexovacímu robotovi ve zjištění všech stránek dané webové aplikace. Hlavní výhodou má na webech s velkým počtem stránek, vyhledávači je tímto souborem usnadněn průchod webovou aplikací.

4.5.20 Vnořené tabulky

Tento test je zařazen spíše z historických důvodů, kdy bylo moderní vytvářet tabulkové vzhledy. Celá stránka byla obsažena v tabulce a někdy se tabulky dokonce vnořovali. Tato technika se velice nedoporučuje a jak bylo zmíněno, tak při nějaké chybě během přenosu se celá stránka nezobrazila. Pokud není ani jedna tabulka na stránce je vypsáno „Žádná tabulka nebyla nalezena“, při výskytu jedné tabulky „Na stránce nejsou vnořené tabulky“, pokud je na stránce více tabulek a nejsou vnořené „Tabulky nejsou vnořené“ a pokud jsou tabulky vnořené „Tabulky jsou vnořené“.

4.5.21 Průchod webem

Tato část testuje dle zadaných vstupních podmínek jednotlivé stránky ve webové aplikaci. Pro tento průchod je možné zvolit způsob a to buď pomocí sitemapy (pokud existuje se standardním názvem) nebo pomocí samovolného průchodu jednotlivými stránkami. Také je možno tento test omezit hloubkou procházení (počtem lomítek v URL) nebo počtem stránek.

Pro získávání zdrojových kódů při průchodu webovou aplikací za použití výše zmíněného způsobu (viz. Obrázek 10) byl problém s překračováním povoleného času pro běh PHP kódu. Proto bylo nutno nasimulovat vlákna a to pomocí *curl* knihovny. Zde je možnost odeslat několik dotazů na servery, kde se nečeká na odpovědi, ale lze je zpracovávat později. Díky čemu lze první přijaté zdrojové kódy otestovat a nemusí se čekat na zdrojové kódy po jednom. Pro příklad: Pokud čas načtení jednoho zdrojového kódu prvním způsobem trval jednu sekundu, tak 20 takovýchto dotazů potrvá 20 sekund. Při použití curl lze čas zkrátit. Nejdůležitějším faktorem je počet povolených přístupů na dotazované aplikaci.

Na každé stránce nejdříve zjistí v hlavičce HTTP odpovědi, zda se jedná o klasický HTML zdrojový dokument a zda je HTTP odpověď s kódem 200.

Pokud není hlavička odpovědi v pořádku, je uživatel informován o tomto problému. Pokud je odpověď v pořádku přechází se k testování jednotlivých částí stránky. Zde není výpis tak obsáhlý, jako u předešlých testů, ovšem jsou vybrány ty nejpodstatnější a jejich výstup je strukturalizován, aby byl přehledný. Také jsou u podstatných nedostatků použity ke zvýraznění barevné mutace. Ke každé testované stránce je navíc připojen odkaz směřující na danou stránku a další pro podrobný test.

4.6 Další funkce SEO aplikace

Pro získávání počtu zpětných odkazů, pozice ve vyhledávačích na určitá slova a celkového počtu zaindexovaných stránek daného webu jsou použita některá API od vyhledávačů. Jelikož výsledky nelze získat jiným způsobem. Jedná se o vyhledávače Google a Bing. Seznam přístupové API nemá. Tyto informace jsou hledány pouze do sté pozice, toto omezení je kvůli šetření dotazů.

4.6.1 Google

Prvním krokem bylo nalezení řešení jak získávat informace od vyhledávače pomocí PHP jazyka [63]. Pro toto řešení je potřeba mít zaregistrovanou službu *Custom Search API* [64], kde jsou k dispozici grafy s využitím a lze zde nastavit limity, základní limit je bezplatný (100 dotazů denně a 10 dotazů za sekundu). Dále je potřeba povolit *OAuth v1* [65] díky čemu je získáno unikátní ID. Poslední potřebnou službou je vytvoření *Vlastního vyhledávání* [66]. Zde je potřeba v nastavení zrušit cíl vyhledávání, díky čemu se vyhledává na celém internetu, nevyplňovat klíčová slova vyhledávače, jelikož ta zužují tematicnost a je zde k nalezení ID vyhledávače, které je potřeba vyplnit v PHP kódu pro přístup. Díky těmto službám je možné vyhledávat na internetu. Právě kvůli limitům pro bezplatné použití je do kódu přidáno pozastavení na dobu 250 milisekund, tím se zamezí porušení limitu i při dvou testech zároveň.

4.6.2 Bing

Pro používání vlastního vyhledávání stačí zaregistrovat a spustit bezplatnou službu [67]. V testovacím prostředí je možné si vybrat cíl vyhledávání, kde je pro tyto účely nejlepší *Web*. A další možnosti, které není potřeba měnit. Dále se vycházelo z PHP kódu [68], ze kterého byl převzat pouze způsob kódování dotazu a dekodování odpovědi.

4.6.3 Seznam

Pro získávání informací od tohoto vyhledávače je použit dotaz velice podobný jako u získání zdrojového kódu testované webové stránky. Poté je stránka parsována a získány potřebné informace. Aby dotazy na vyhledávač nebyly moc rychlé, je zde také použito uspání na dobu 250 milisekund.

4.6.4 Zpětné odkazy

Jak bylo zmíněno v teoretické části, tak mezi hlavní Off-page faktory patří zpětné odkazy. Pro tyto účely byla vytvořena samostatná stránka, která zjišťuje počet zpětných odkazů mířících na danou stránku či doménu podle možností vyhledávače. Uživatel si může vybrat, které vyhledávače si přeje zobrazit ve výsledcích. Tento test je určen pouze jako informativní, aby uživatel viděl, zda jeho práce na linkbuildingu se vyplácí či je potřeba zvolit jiný směr budování zpětných odkazů.

4.6.5 Pozice ve vyhledávačích

Na této stránce jak napovídá nadpis lze snadno kontrolovat pozice webové stránky na určitá klíčová slova, které si uživatel zadá sám. I u tohoto testu lze vybrat, pro jaké vyhledávače chceme zobrazit výsledky. Test vypíše první nalezenou pozici, pokud tato pozice je menší nebo rovna než stý výsledek ve vyhledávání. Také zobrazuje celkový počet nalezených výsledků. Další informaci, kterou zobrazí je počet zaindexovaných stránek u daného vyhledávače.

4.6.6 Ranky

Tato stránka slouží pouze pro zjištění hodnot S-Ranku, PageRanku a jako kurziozítka Alexa Ranku. První dva ranky jsou pro české uživatele celkem podstatné. Bing do této kategorie není zařazen, jelikož nezveřejňuje hodnotu ranku, na místo toho byl právě přidán Alexa Rank. I u tohoto testu je možno si vybrat, které ranky chce uživatel zobrazit.

4.6.7 Analýza validity

Tato část je v aplikaci hlavně pro to, aby uživatelé viděli správnost a kvalitu svého kódu. Další a důležitější důvod tohoto testu je pokud nelze vytvořit SEO analýzu, nebo pokud je ve výsledcích hodně nesrovnalostí s originálním zdrojovým kódem, tak chyba bude nejspíše ve velkém počtu validačních chyb ve zdrojovém kódu. Je zde přidána i validace CSS stylů, která by měla zaručit správné vykreslování prvků napříč webovými prohlížeči.

4.7 Jednotná informační vrstva

V celé aplikaci je díky frameworkům použito několik informačních systémů. Jedním z nich jsou informace z formulářů, a to o úspěšnosti odeslání dotazu na server. Dalším je informace o probíhajících testech, které jsou aktivovány pomocí AJAX technologie. A poslední jednotnou vrstvou jsou informace k jednotlivým testům.

4.8 Registrovaná část

V této části jsou naprosto shodné testy, ovšem odpadá nutnost vkládání cílové adresy webové stránky. Registrovaná část začíná registrací (viz. Obrázek 11). Prvním ošetřením je shoda hesel, která je ošetřena na straně prohlížeče pomocí

JavaScriptu, toto ošetření je volitelnou částí frameworku. Po odeslání je zkontrolována emailová adresa s adresami v databázi, pokud není adresa nalezena, tak jsou údaje uloženy a heslo je zašifrováno jednosměrnou šifrou kvůli bezpečnosti a aby se tento údaj nemohl zneužít provozovatelem a vlastníkem. Po uložení údajů do databáze je uživatel automaticky přihlášen. A přesměrován do registrované části do správy aplikací. Toto přesměrování nastane i po přihlášení (viz. Obrázek 12), ovšem pokud je již uživatel registrován.

Při přihlášení do aplikace a nezaškrtnutí Zapamatovat přihlášení je uživatel odhlášen po 10 minutách nečinnosti, nebo při uzavření prohlížeče. Pokud je políčko zaškrtnuto, přihlášení je aktivní 14 dní z daného počítače a prohlížeče, při každé návštěvě se tato doba opět nastaví na zmíněnou dobu.

Registrace

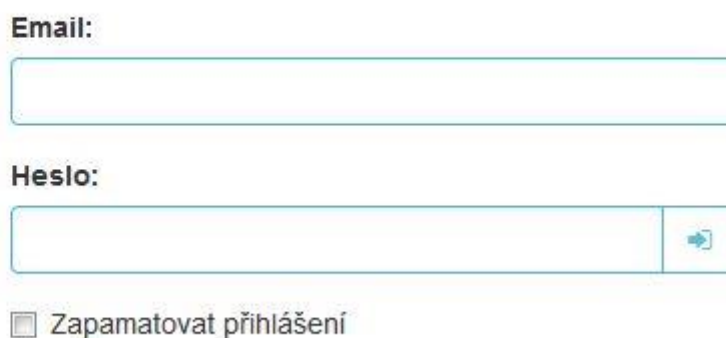


Registration form with the following fields:


- Jméno:
- Příjmení:
- Email:
- Vaše heslo:
- Heslo znovu: 

Obrázek 11: Registrace

Přihlášení



Login form with the following fields:

- Email:
- Heslo: 
- ☐ Zapamatovat přihlášení

Obrázek 12: Přihlášení

4.8.1 Správa aplikací

Na této stránce je možnost přidávat a odebírat sledované webové aplikace. Aplikace se přidávají na pravé části stránky bez protokolu `http://` a adresa musí směřovat na index. V prostřední části stránky jsou vypsané jednotlivé přidané aplikace uživatele s možností odebrání (viz. Obrázek 13).

Po vybrání aplikace je zobrazen výpis s některými posledními hodnotami z testů a jsou součástí i dobře měřitelné informace o zlepšení (viz. Obrázek 14). V pravé části této stránky je možnost zařazení do diskuze pro rozumnou výměnu zpětných odkazů. Tato možnost je plně volitelná a lze ji kdykoliv opět zrušit.

Webové aplikace

The screenshot displays a web application management interface. On the left, under the heading "Vaše weby" (Your websites), there is a list of five tracked websites, each in a light blue box with an "Odeber" (Remove) button to its right:

- tui.cz
- alza.cz
- cybershopcz.com
- web.navyseal.cz
- realitysluknov.cz

On the right, under the heading "Přidat webovou aplikaci" (Add web application), there is a form titled "Vložte adresu webu na index" (Paste the website address on the index). It contains a text input field with the placeholder "Adresa bez http://" and a small blue icon button to its right.

Obrázek 13: Správa aplikací

Webová aplikace

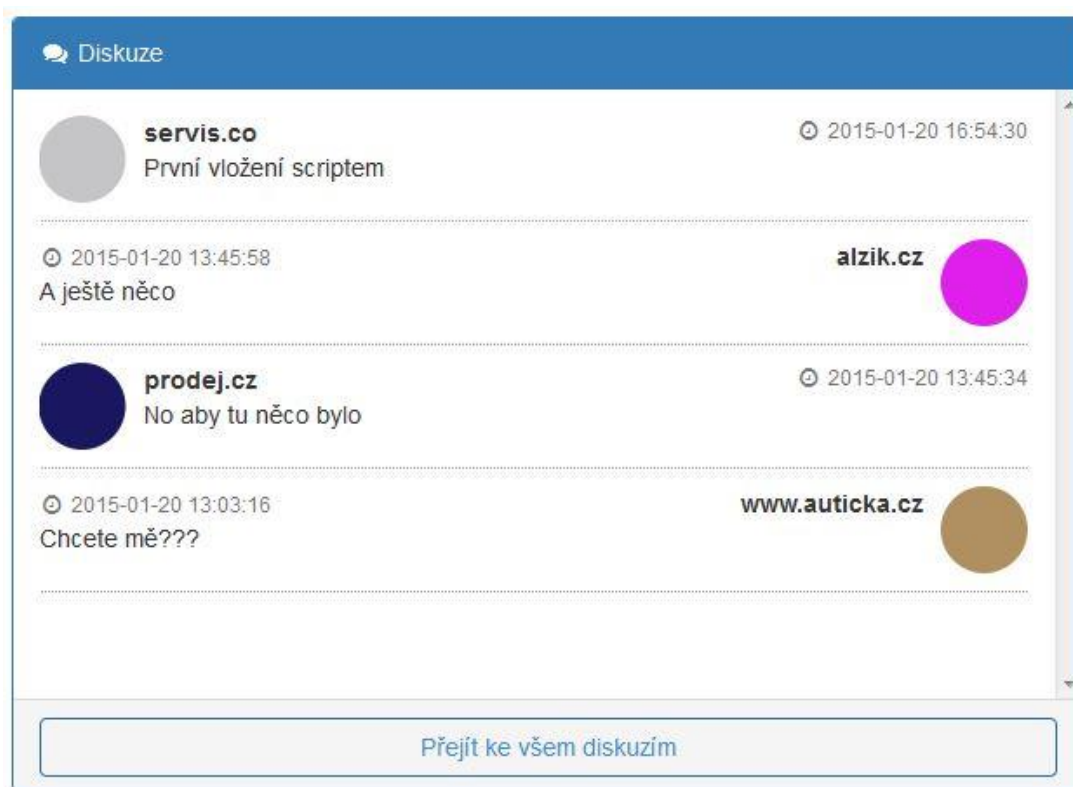
The screenshot displays a web application interface. On the left, a green header bar contains the URL 'cybershopcz.com'. Below it, a list of statistics is shown: 'Web přidán: 26. 01. 2015', 'Testů na počet zaindexovaných stránek: 1', 'Testů seo: 2', 'Testů pozice na výraz: 1', 'Testů ranky: 1', and 'Testů na zpětné odkazy: 1'. A section titled 'Poslední testy' follows, listing various metrics such as 'Počet zaindexovaných stránek: 265', 'Čas načtení zdrojového kódu: 591 microsekund', 'Délka zdrojového kódu: 35846 znaků', 'Počet slov na stránce: 8411', 'Počet odkazů: 57', 'Počet odkazů ven: 1', 'Zlepšení: PageRank: 0, S-Rank: 0', 'Počet zpětných odkazů: 16', and 'Poslední test: 20. 04. 2015'. On the right, a blue header bar is labeled 'Nastavení tématu webu a diskuze'. Below it, there is a checkbox labeled 'Zahrnout tento web do nabídky zpětných odkazů' which is currently unchecked, and a button labeled 'Uložit'.

Obrázek 14: Vybraná webová aplikace

Po zaškrtnutí zahrnutí aplikace do diskuze se zpětnými odkazy je zobrazena nabídka, která obsahuje tematické okruhy a jejich upřesnění. Po vybrání tematického okruhu se pomocí AJAXu načtou upřesňující kategorie (viz. Obrázek 15). Po uložení těchto informací dojde k vykreslení malého okénka s diskuzemi na dané téma (viz. Obrázek 16).

This screenshot shows the 'Nastavení tématu webu a diskuze' settings panel. At the top, the title 'Nastavení tématu webu a diskuze' is displayed. Below it, a checkbox labeled 'Zahrnout tento web do nabídky zpětných odkazů' is checked. A section titled 'Prosím vyberte tematický obor Vaší stránky' contains two dropdown menus. The first, labeled 'Kategorie', has 'Auto - moto' selected. The second, labeled 'Podkategorie tématu', has 'Prodej' selected. At the bottom left of the panel is a button labeled 'Uložit'.

Obrázek 15: Nastavení tématu diskuze



Obrázek 16: Miniaturní diskuze

4.8.2 Diskuze

Pokud si uživatel není naprosto jist přesným zaměřením stránky, tak může přejít ke všem diskuzím na dané téma. A zde si zvolit buď jednotlivá zaměření, nebo si nechat vypsat všechna zaměření. Stejným způsobem lze i vkládat nové diskuze (viz. Obrázek 17). Po kliknutí na diskuzi se dostanete na chatovací stránku (viz. Obrázek 18), která se sama aktualizuje každých 10 sekund. Jednotlivé položky v diskuzi i chatu mají pro zpříjemnění barevné puntíky, kde se barvy náhodně generují.

Diskuze -> Auto - moto

Filtr podkategorie

Servis

auticka.cz

Sháníme partnery

Servis

← Zpět

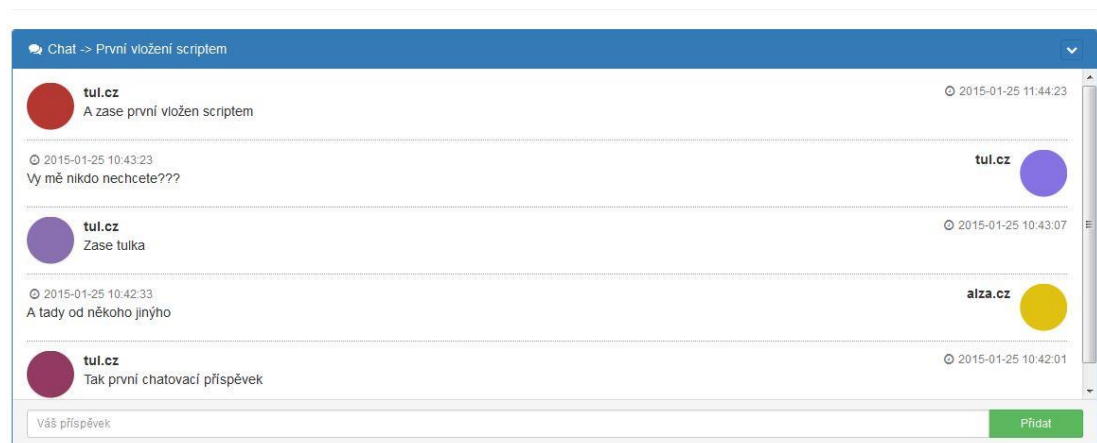
© 2015-01-20 13:02:30

Název Vaší diskuze

Prodej

Vytvořit

Chat



57

5 Vyhodnocení řešení

5.1 Časová a paměťová náročnost

Na webové aplikace se klade celkem velký důraz na rychlé vyhodnocení a sestavení výsledného zdrojového kódu s co nejmenší paměťovou náročností. To je způsobeno mnohými omezeními na webových hostinzích.

Byly vybrány čtyři webové aplikace, které mají rozdílně velký zdrojový kód a nemají společný server. V následující tabulce je přehled analyzovaných stránek včetně velikosti každé stránky a času načtení. Velikost stránky je velikost zdrojového kódu stránky na příslušné adrese bez externích souborů, jako jsou kaskádové styly, obrázky apod.

Tabulka 1: Velikost a čas načtení webových stránek

adresa webové stránky	velikost	čas načtení zdrojového kódu
http://dentic.cz	18 555 B	262 ms
http://realitysluknov.cz	13 403 B	140 ms
http://alza.cz	177 805 B	342 ms
http://heureka.cz	40 386 B	124 ms

Testování bylo provedeno ve vývojovém prostředí, rozdíl mezi vývojovým a produkčním režimem je zanedbatelný. Testován byl čas a paměť, za který se provedla analýza a načtení zdrojového kódu do aplikace. Alokovaná paměť je část paměti, kterou si vyhradila aplikace pro svoje výpočty a meziukládání. Díky *Tracy* jsou zmíněné informace součástí výsledku aplikace ve vývojovém režimu.

Tabulka 2: Časová a paměťová náročnost - domácí počítač

adresa webové stránky	čas analýzy	využitá paměť
http://dentic.cz	521 ms	6,53 MB
http://realitysluknov.cz	546 ms	6,55 MB
http://alza.cz	1 068 ms	6,62 MB
http://heureka.cz	837 ms	6,73 MB

Tabulka 3: Časová a paměťová náročnost - server hosting

adresa webové stránky	čas analýzy	využitá paměť
http://dentic.cz	452 ms	3,66 MB
http://realitysluknov.cz	498 ms	3,66 MB
http://alza.cz	1 030 ms	4,87 MB
http://heureka.cz	737 ms	3,86 MB

5.2 Zlepšení pozic testované webové aplikace

Tato aplikace byla během svého vývoje nasazena na webhosting s adresou <http://seopomoc.lukasjavurek.eu>. Začátkem vývoje byly metainformace v hlavičce naplněny krátkými a někde i duplicitními informacemi. Po téměř dokončeném vývoji došlo k analýze „sebe sama“ (stránky aplikace byli otestovány aplikací). Na základě doporučení byly provedeny změny v šabloně. Tyto změny vedli k posunutí ve výsledcích ve vyhledávání. Posunutí činilo zhruba o dvě až čtyři stránky kupředu. Nyní se aplikace zobrazuje od první stránky do čtvrté, samozřejmě i dále, jelikož na hodně obecná a užívaná slova nemá podporu zpětných odkazů.

Testované fráze v anonymním režimu prohlížeče, který nikdy nebyl na dané adrese:

- seo pomoc
- seo pomoc testy
- seo analýza pomoc informace
- seo testy analýza

5.3 Srovnání existujících aplikací

Pro srovnání s podobnými webovými aplikacemi byly vybrány ty aplikace, které se umísťují na předních pozicích ve vyhledávačích. A byly vybrány testy zaměřené na zdrojový kód.

5.3.1 Seo-servis

Seo-servis je webový portál zabývající se SEO problematikou. Obsahuje sadu nástrojů pro optimalizaci webových stránek pro vyhledávače. Zde za poslední rok rozšířili testy a doplnili doprovodné informace z vyhledávačů. Portál je rozdělen na následující služby:

- zdrojový kód
- klíčová slova
- vyhledávače
- síla webu
- zpětné odkazy

Ovšem chybí zde doplňující informace k jednotlivým testům, jako je například důvod výskytu. A u jednotlivých výsledků je poměrně malé vysvětlení výsledných informací. Toto je způsobeno tím, že nabízejí placené školení na téma SEO, kde by měli vysvětlit jednotlivé výsledky.

5.3.2 Seonastroje

Seonastroje je poměrně komplexní webová aplikace. Tato aplikace je v celku dobře členěná na tematické skupinky. Tato aplikace se také za poslední rok velice rozšířila a přidala funkcionality. Rozdělení je následující:

- SEO Monitor
- SEO Analýza
- SEO Analýza – srovnání
- Analýza klíčových slov
- Pozice ve vyhledávačích
- Kontrola zpětných odkazů
- Validace HTML a CSS
- Ranky, indexy...
- Hledanost frází
- Odkazy na doménu

Sice na první pohled může vypadat, že aplikace toho umí opravdu hodně, ale po důkladnějším prozkoumání se výsledky poměrně opakují a jsou použity vyhledávače, které v dnešní době čeští uživatelé nevyužívají a ani v historii nebyly moc využívány. Část, která stojí za zmínku je SEO Monitor, který sjednocuje monitorování pozic ve vyhledávačích, kontroluje ranky a tyto změny zasílá na email uživateli. Toto monitorování dělá pravidelně. A také podporuje zakázanou techniku SEO, kterou je kupování zpětných odkazů. V této aplikaci také chybí vysvětlení výsledných informací u testů.

5.3.3 Otestujte Web

Otestujte Web je aplikace, která oproti ostatním má pouze jednu stránku s testy. Tento test je celkem komplexní. Po analýze se zobrazí Celkové hodnocení se zkrácenými výsledky. Dále si uživatel může prohlédnout podrobnější výpisy:

- Hlavička
- Soubory pro spolupráci s roboty a vyhledávači
- Struktura
- Zpětné odkazy
- Síla webu
- Socializace

V této aplikaci stojí za zmínku kategorie Socializace, zde jsou prohledávány sociální sítě jako Facebook, Twitter, Google+ a zjišťuje počty zpětných odkazů na testované stránky. Ovšem jako u předešlých aplikací i zde chybějí nějaké informace k výsledkům.

Tabulka 4: Srovnání rychlosti testu na zdrojovém kódu

SEO nástroj	čas testu realitysluknov.cz
http://otestujteweb.cz	5 010 ms
http://seonastroje.cz	1 530 ms
http://seo-servis.cz	2 110 ms
tento nástroj	546 ms

5.4 Možnosti rozšíření

Díky objektovému přístupu v celé aplikaci není velký problém tyto testy rozšiřovat a přidávat nové funkcionality.

Tato práce je zaměřena převážně na On-page faktory (SEO analýza), ovšem díky testům na Ranky a Zpětné odkazy je zaměřena i na Off-page faktory. Výpis testu na zpětné odkazy by se mohl rozšířit. Vypisovat počty odkazů ze sociálních sítí, signály ze sociálních sítí (like, +1) a odkazy z jiných stránek.

Další možností je zaměření se na zakázané metody, za které dávají internetové vyhledávače penalizace popřípadě ban, a tudíž ovlivňují výsledek ve vyhledávání.

Možným rozšířením by bylo vytvořit JavaScript, který by zjišťoval návštěvnost na sledované webové aplikaci a ukládal statistiky přístupu a pohybu uživatelů na stránkách.

Všechna tato rozšíření by byla poměrně náročná na server, tudíž aplikace by měla být na vlastním serveru (nikoli na sdíleném webhostingu, kde může být problém s nadměrnou komunikací s ostatními webovými aplikacemi). Dalším problémem by byla nutnost placených API přístupů k vyhledávačům, jelikož bezplatné verze mají příliš malý počet dotazů na vyhledávače.

6 Závěr

Hlavním cílem práce bylo stanovit testovatelné části webových aplikací, které jsou důležité z hlediska SEO. Jelikož se informace rozcházejí v důležitosti výskytu jednotlivých částí webové stránky, bylo potřeba tuto problematiku dopodrobna nastudovat a vyzkoušet nabitě znalosti přímo ve fulltextových vyhledávačích. V teoretické části jsou tyto faktory seskupené do celku, který podává komplexní přehled o této problematice a jejích správných postupech.

Na základě získaných informací byla vytvořena webová aplikace, která umožňuje testovat jakékoliv webové stránky a je rozdělena na tematicky podobné stránky s testy. Díky čemuž jsou testy rychlé, tematicky seskupené a nedochází ke zbytečným testům ve vyhledávačích.

Jedno z hlavních pozitiv této aplikace je absence bodové či procentuální úspěšnosti ve výsledcích testů a existence rozsáhlé informační vrstvy, která obsahuje rady a tipy na zlepšení daných částí webové stránky, jelikož takovýto komplexnější systém není v české lokaci.

Aplikace je určena převážně amatérům, začátečníkům, středně pokročilým a některé funkce jsou vhodné i pro profesionály. Informační vrstva je určena hlavně začátečníkům, jelikož je psána srozumitelnou a jednoduchou formou, kde je text poměrně strukturalizován, a pro rozšíření informací jsou připojeny odkazy na komplexnější články, jelikož hlavním cílem nebyla komplexní informační vrstva, ale pouze rady s tipy na zlepšení daných částí.

Webová aplikace musí využívat některá API od vyhledávačů, jelikož nedovolují strojové dotazy a umí poměrně dobře odhalit přístup pomocí scriptů. Google nabízí API pro vyhledávání, kde starší verze je stabilní, ale s chybami, u jedné části musel být upraven zdrojový kód, jelikož velice často vyvolávala výjimku. Další nedokonalostí API je, že některé jednoslovné výrazy nevyhledává, místo toho vygeneruje výjimku. Řešením by mohlo být dvojité zadání hledaného slova, ovšem toto ovlivňuje výsledky. Existuje nová verze, tato verze je ovšem stále beta verzí (kontrolováno 12. 5. 2015).

Výsledná webová aplikace je snadno rozšiřitelná díky použitému frameworku a objektům. Aplikace by mohla být rozšířena o testy zaměřené na zakázané techniky

SEO. Dalším rozšířením by mohlo být zjišťování, zda doména získala penalizace od vyhledávačů.

Použité zdroje

- [1] *Internetový katalog – Wikipedie* [online]. [vid. 21. 11. 2013].
Dostupné z: http://cs.wikipedia.org/wiki/Internetový_katalog
- [2] Janovský, Dušan. *Katalogy* [online].
Dostupné z: <http://www.jakpsatweb.cz/katalogy.html>
- [3] *Fulltextové vyhledávání – Wikipedie* [online]. [vid. 24. 11. 2014].
Dostupné z: http://cs.wikipedia.org/wiki/Fulltextové_vyhledávání
- [4] Janovský, Dušan. *Vyhledávače, jak fungují* [online]. [vid. 2012].
Dostupné z: <http://www.jakpsatweb.cz/vyhledavace.html>
- [5] *Google (vyhledávač) – Wikipedie* [online]. [vid. 11. 2. 2015].
Dostupné z: [http://cs.wikipedia.org/wiki/Google_\(vyhledávač\)](http://cs.wikipedia.org/wiki/Google_(vyhledávač))
- [6] *Hlavní strana – SEO* [online].
Dostupné z: <http://www.seo-optimalizace.info/>
- [7] *Seo – Wikipedia* [online]. [vid. 11. 2. 2015].
Dostupné z: <http://cs.wikipedia.org/wiki/Seo>
- [8] Tichý, Jan. *Úvod do praktického SEO* [online]. [vid. 23. 4. 2013].
Dostupné z: <https://youtu.be/wNQVwWQfYHw>
- [9] *Než začnete tvořit web | Seznam Nápověda* [online].
Dostupné z: <http://napoveda.seznam.cz/cz/fulltext-hledani-v-internetu/nez-zacnete-tvorit-web/>
- [10] *On-page faktory | Seznam Nápověda* [online].
Dostupné z: <http://napoveda.seznam.cz/cz/fulltext-hledani-v-internetu/optimalizace-webu/on-page-factory/>
- [11] Janovský, Dušan. *Duplicity a podobnosti* [online].
Dostupné z: <http://www.jakpsatweb.cz/seo/duplicity-podobnosti.html>
- [12] *Duplicitní obsah – Wikipedie* [online]. [vid. 9. 6. 2013].
Dostupné z: http://cs.wikipedia.org/wiki/Duplicitní_obsah

- [13] Janovský, Dušan. *Přirozené duplicity a kanonizace* [online].
Dostupné z: <http://www.jakpsatweb.cz/seo/kanonizace.html>
- [14] Kylián, Ivo. *Identifikace problémů s načítáním webu - IT Logica* [online].
Dostupné z: <http://blog.it-logica.cz/jak-identifikovat-rychlost-stranek>
- [15] *Nejčastější chyby při linkbuildingu | Comerto - tvorba webu, internetový marketing, SEO* [online]. [vid. 19. 11. 2012].
Dostupné z: <http://www.comerto.com/poradenstvi-skoleni/odborne-clanky/nejcastejsi-chyby-pri-linkbuildingu>
- [16] Janovský, Dušan. *Optimální tvar URL* [online].
Dostupné z: <http://www.jakpsatweb.cz/seo/seo-url.html>
- [17] Janovský, Dušan. *Doména pro SEO* [online]. [vid. 2007].
Dostupné z: <http://www.jakpsatweb.cz/seo/seo-domena.html>
- [18] Janovský, Dušan. *DTD - Document type definition v HTML* [online].
Dostupné z: <http://www.jakpsatweb.cz/doctype.html>
- [19] *HTML head Elements – w3schools.com* [online].
Dostupné z: http://www.w3schools.com/html/html_head.asp
- [20] Facebook.com. *The Open Graph protocol* [online].
Dostupné z: <http://opengraphprotocol.org/>
- [21] *On-page faktory | Seznam Nápořěda* [online].
Dostupné z: <http://napoveda.seznam.cz/cz/fulltext-hledani-v-internetu/optimalizace-webu/on-page-faktory/>
- [22] Havrlant, Lukáš. *Sémantika - Úvod do sémantiky* [online].
Dostupné z: <http://www.semantika.name/uvod-semantika.html>
- [23] *Off-page faktory | Seznam Nápořěda* [online].
Dostupné z: <http://napoveda.seznam.cz/cz/off-page-faktory.html>
- [24] *Offpage faktory – Wikipedie* [online]. [vid. 14. 9. 2009].
Dostupné z: http://cs.wikipedia.org/wiki/Offpage_faktory

- [25] *PageRank – Wikipedie* [online]. [vid. 11. 10. 2014].
Dostupné z: <http://cs.wikipedia.org/wiki/PageRank>
- [26] *S-Rank – Wikipedie* [online]. [vid. 11. 5. 2014].
Dostupné z: <http://cs.wikipedia.org/wiki/S-Rank>
- [27] Janovský, Dušan. *Google PageRank, vysvětlení a odpovědi* [online].
Dostupné z: <http://www.jakpsatweb.cz/seo/pagerank.html>
- [28] *S-rank (SRank) ukazatel* [online].
Dostupné z: <http://www.webhosting-domeny.info/seo/s-rank.php>
- [29] Internet-studio. *Zpětné odkazy - Srank, Pagerank, Alexarank* [online].
Dostupné z: <http://www.zpetne-odkazy.info/srank-pagerank-alexarank.html>
- [30] Analogic. *Ranky - PageRank, SRank, Jyxo Rank, Alexa* [online].
Dostupné z: <http://www.webhosting-domeny.info/seo/s-rank.php>
- [31] *S-rank | Seznam Nápořěda* [online].
<http://napoveda.seznam.cz/cz/fulltext-hledani-v-internetu/hledani-fulltext-algoritmus-vyhledavani-razeni-vysledku-faq-dotazy/>
- [32] *Alexa Internet - Toolbar Download* [online].
<http://www.alexa.com/toolbar>
- [33] *Alexa rank / SEO-Expert.cz* [online].
<http://www.seo-expert.cz/alexa-rank>
- [34] *Alexa Internet – Wikipedie* [online]. [vid. 7. 3. 2015].
http://cs.wikipedia.org/wiki/Alexa_Internet
- [35] *Black hat SEO / SEO-Expert.cz* [online].
<http://www.seo-expert.cz/black-hat-seo>
- [36] *Black Hat SEO / SEO - Profesionál.cz* [online].
<http://www.seo-profesional.cz/black-hat-seo/>

- [37] *Zakázané SEO praktiky optimalizace webových stránek | Propagace na internetu* [online]. [vid. 29. 4. 2011].
Dostupné z: <http://www.propagacenainternetu.cz/zakazane-seo-praktiky>
- [38] Dočekal, Daniel. *SEO slovník* [online].
Dostupné z: <http://www.effectix.com/cz/blog/seo-slovník>
- [39] Hodboď, Tomáš. *SEO spam: zneužití umění* [online]. [vid. 25. 4. 2005]
Dostupné z: <http://www.lupa.cz/clanky/seo-spam-zneužite-umeni/>
- [40] *SEO slovník | SEO - Profesionál.cz* [online].
Dostupné z: <http://www.seo-profesional.cz/seo-slovník/>
- [41] *Zakázané optimalizační techniky | Seznam Nápořěda* [online].
Dostupné z: <http://napoveda.seznam.cz/cz/fulltext-hledani-v-internetu/ceho-se-pri-optimalizaci-vyvarovat/zakazane-optimalizacni-techniky/>
- [42] Gigant GROUP s.r.o. *SEO spam - Zakázané praktiky SEO optimalizace | ALTOS.cz* [online]. Dostupné z: <http://seo-altos.altos.cz/seo-spam/>
- [43] *Komentářový spam = diskusní spam | SEO-Expert.cz* [online].
Dostupné z: <http://www.seo-expert.cz/komentarovy-spam-diskusni-spam>
- [44] *Cloaking | SEO-Expert.cz* [online].
Dostupné z: <http://www.seo-expert.cz/cloaking>
- [45] *Doorway page | SEO-Expert.cz* [online].
Dostupné z: <http://www.seo-expert.cz/doorway-page>
- [46] *Link farm | SEO-Expert.cz* [online].
Dostupné z: <http://www.seo-expert.cz/link-farm>
- [47] *Blog SEO-PPC.cz: Microsite a SEO* [online].
Dostupné z: <http://blog.seo-laboratory.cz/p/microsite-seo.html>
- [48] *Nette Framework – Wikipedie* [online]. [vid. 8. 4. 2015]
Dostupné z: http://cs.wikipedia.org/wiki/Nette_Framework

- [49] Skvorc, Bruno. *Best PHP Framework for 2015 - SitePoint Survey Results* [online]. [vid. 28. 3. 2015]. Dostupné z: <http://www.sitepoint.com/best-php-framework-2015-sitepoint-survey-results/>
- [50] Nette Foundation. *Rychlý a pohodlný vývoj webových aplikací v PHP* [online]. [vid. 29. 3. 2015] Dostupné z: <http://nette.org/cs/#toc-features>
- [51] Nette Foundation. *Požadavky Nette Framework* [online]. [vid. 6. 8. 2014]. Dostupné z: <http://doc.nette.org/cs/2.2/requirements>
- [52] Nette Foundation. *Začínáme* [online]. [vid. 18. 3. 2015]. Dostupné z: <http://doc.nette.org/cs/2.3/quickstart/getting-started>
- [53] Nette Foundation. *Debugování a zpracování chyb* [online]. [vid. 5. 3. 2015]. Dostupné z: <http://doc.nette.org/cs/2.3/debugging>
- [54] Nette Foundation. *Šablony* [online]. [vid. 12. 3. 2015]. Dostupné z: <http://doc.nette.org/cs/2.3/templating>
- [55] Nette Foundation. *MVC aplikace & presentery* [online]. [vid. 2. 4. 2015]. Dostupné z: <http://doc.nette.org/cs/2.3/presenters>
- [56] Nette Foundation. *Databáze* [online]. [vid. 21. 1. 2015]. Dostupné z: <http://doc.nette.org/cs/2.3/database>
- [57] Twitter Bootstrap. *bootstrap/LICENSE at master* [online]. [vid. 1. 1. 2015]. Dostupné z: <https://github.com/twbs/bootstrap/blob/master/LICENSE>
- [58] *Twitter Bootstrap – Wikipedie* [online]. [vid. 12. 11. 2014]. Dostupné z: http://cs.wikipedia.org/wiki/Twitter_Bootstrap
- [59] *GNU General Public License – Wikipedie* [online]. [vid. 27. 10. 2014]. Dostupné z: http://cs.wikipedia.org/wiki/GNU_General_Public_License
- [60] *MySQL – Wikipedie* [online]. [vid. 3. 1. 2015] Dostupné z: <http://cs.wikipedia.org/wiki/MySQL>
- [61] *MyISAM – Wikipedie* [online]. [vid. 9. 8. 2014]. Dostupné z: <http://cs.wikipedia.org/wiki/MyISAM>

- [62] *InnoDB – Wikipedie* [online]. [vid. 11. 12. 2014].
Dostupné z: <http://cs.wikipedia.org/wiki/InnoDB>
- [63] Melrose, Stephen. *Google APIs PHP clients* [online]. [vid. 6. 1. 2012].
Dostupné z: <https://github.com/stephenmelrose/google-apis-php-clients/tree/fc4f12a7df8fce13351ec3b1254a8b288ab24ce8>
- [64] Google. *Custom Search API* [online]. Dostupné z:
<https://console.developers.google.com/project/271601827204/apiui/apiview/customsearch/overview>
- [65] Google. *OAuth v1* [online]. Dostupné z:
<https://console.developers.google.com/project/271601827204/apiui/credential>
- [66] Google. *Vlastní vyhledávání* [online].
Dostupné z: <https://cse.google.cz/cse/all>
- [67] Bing. *Developer Resources* [online].
Dostupné z: <http://www.bing.com/dev/en-us/dev-center>
- [68] Tian. *Bing search* [online]. Dostupné z:
https://github.com/songty/bing-search/blob/master/bing_proxy.php

A Obsah přiloženého CD

- text_bakalarske_prace

Adresář obsahuje veškerý text bakalářské práce.

- Bakalarska_prace_2015_Lukas_Javurek.pdf
- Kopie_zadani_bakalarske_prace_2015_Lukas_Javurek.pdf

- zdrojove_soubory_aplikace

V adresáři jsou obsaženy kompletní zdrojové kódy výsledné aplikace.

- obrazky_v_bakalarske_praci

Adresář obsahuje obrázky použité v textové části bakalářské práce.

- snimky_obrazovky_vysledne_aplikace

Adresář obsahuje snímky z veškerých hlavních částí výsledné aplikace.